



**Aalto-yliopisto**  
Insinööritieteiden  
korkeakoulu

Lasse Kangas

## **Hyperspektrisen superjatkumo-lidarin käyttö kiviaineksen tutkimisessa**

Diplomityö, joka on jätetty opinnäytteenä tarkastettavaksi diplomi-insinöörin tutkintoa varten.

Espoossa 25.2.2016

Valvoja ja ohjaaja: Professori Jussi Leveinen

---

**Tekijä** Lasse Kangas

---

**Työn nimi** Hyperspektrisen superjatkumo-lidarin käyttö kiviaineksen tutkimisessa

---

**Koulutusohjelma** Yhdyskunta- ja ympäristötekniikka

---

**Pää-/sivuaine** Teknillinen geologia ja sovellettu geofysiikka

**Koodi** IA3029

---

**Työn valvoja** Prof. Jussi Leveinen

---

**Työn ohjaaja(t)** Prof. Jussi Leveinen

---

**Päivämäärä** 19.2.2016

**Sivumäärä** 72

**Kieli** Suomi

---

Kiviaineshuollossa käytetyn kiviaineksen laadun arvioimiseen käytetyt menetelmät ovat perinteisesti vaatineet geologien kenttäkatselmuksia ja laboratoriossa standardien mukaisesti suoritettuja tutkimuksia. Kiviaineksen oton aikainen laadunvarmistus on useampien laatuparametrien osalta hyvin harvoin toistuvaa, joten louhinnan tai oton edistyessä kiviaineksen laatumuutokset saattavat jäädä huomaamatta tai ne huomataan liian myöhään. Tämä saattaa johtaa tilanteeseen, jossa lopputuotteen käyttöikä ei ole suunnitellun mukainen. Perinteisten laadunhallinnan työkalujen rinnalle on nousemassa automatisoituja laadunvarmistusmenetelmiä, joiden etuna on reaaliaikainen *in-situ* tai *on-site* mittaus.

Tässä diplomityössä tutkittiin hyperspektrisen superjatkumo-lidarin käyttökelpoisuutta kiviainesten tutkimisessa. Näytteinä oli yleisesti käytössä olevia rakennuskiviä, jotka pyrittiin erottamaan toisistaan spektrisen sormenjäljen perusteella. Mitatun spektritiedon hyödyntäminen perustuu erilaisten luokittelualgoritmien käyttöön ja edustavien referenssi-luokkien määrittämiseen. Käytetyt luokittelumenetelmät olivat spektrikulman mukainen luokittelu, etäisyysluokittelu, korrelaatioluokittelu, kanoninen erotteluanalyysi, ryhmittelyanalyysi ja pääkomponenttianalyysi.

Spektrikulman ja korrelaation avulla tehdyt analyysit antoivat heikoimmat tulokset. Etäisyysluokittelu ja ryhmittelyanalyysi luokittelivat molemmat kiviainekset yli 95 prosentin tarkkuudella oikeisiin luokkiin. Kanonisen erotteluanalyysin avulla saatiin parhaat tulokset, kokonaisluokittelutarkkuuden ollessa yli 99 prosenttia. Pääkomponenttianalyysia käytettiin dimensioiden vähentämisen vaikutusten tutkimiseen. Muuttujien määrän vähentäminen kahdeksasta alkuperäisestä muuttujasta kolmeen pääkomponenttiin heikensi ryhmittelyanalyysin tuottamaan luokittelutarkkuutta vain noin yhdellä prosenttiyksiköllä.

Saadut tulokset ovat rohkaisevia ja viittaavat siihen, että uudet menetelmät voisivat olla tulevaisuudessa osa kiviainesten laadunvalvontaa mahdollistaen samalla yhä reaaliaikaisemman seurannan. Myös olemassa oleva tutkimustieto viittaa siihen, että hyperspektritiedon ja kiviainesten laatuparametrien välillä on selkeä yhteys.

---

**Avainsanat:** Kiviaines, tunnistaminen, hyperspektri, luokittelu, superjatkumo, lidar

---



---

**Author** Lasse Kangas

---

**Title of thesis** Identification of rocks using a hyperspectral supercontinuum lidar

---

**Degree programme** Civil and Environmental Engineering

---

**Major/minor** Engineering Geology and Applied Geophysics

**Code** IA3029

---

**Thesis supervisor** Prof. Jussi Leveinen

---

**Thesis advisor(s)** Prof. Jussi Leveinen

---

**Date** 19.2.2016

**Number of pages** 72

**Language** Finnish

---

## **Abstract**

The methods used to define the physical quality of aggregates are traditionally based on field research and on standardized laboratory experiments. As a result of these time consuming processes, the quality controls are quite rarely repeated. When the extraction in the quarry proceeds, there might be some changes in the quality of the aggregates that are unnoticed. This might lead to a case where the end product is not what it should be and the lifetime is much shorter than expected. Therefore, automated real-time *in-situ* or *on-site* methods are developed to complement the traditional quality control methods.

In this Master's Thesis, the usefulness of a hyperspectral supercontinuum lidar as a tool for rock type definition was examined. The samples that were used were common building stones. The idea was to discriminate the samples based on their spectral fingerprint. The utilization of the spectral data is based on different classification algorithms and to the determination of representative reference categories. The methods used were: Spectral angle mapper, distance classification, spectral correlation mapper, canonical discriminant analysis, cluster analysis and principal component analysis.

The spectral angle mapper and the spectral correlation mapper gave the weakest results. The distance classification and the cluster analysis were both able to classify the rocks with over 95 percent accuracy. The canonical discriminant analysis gave the best results with a classification accuracy of over 99 percent. The principal component transformation was used to exam the effects of dimension reduction of the data. The reduction from eight original variables to three principal components reduced the accuracy of the cluster analysis only with less than one percentage point.

The results gain from this study were encouraging and they suggest that in the future, these new methods could be a part of real-time rock quality control. In addition, the existing research data indicates that there is a connection between the quality parameters of rock and their hyperspectral data.

---

**Keywords** Rocks, aggregates, classification, hyperspectral, supercontinuum, lidar

---

## Alkusanat

*Tämä diplomityö tehtiin osana Etelä-Suomen VEDET – kärkihanketta. Työ tehtiin Uudenmaan liiton toimeksiantona ja rahoituksella.*

*Haluan kiittää Uudenmaan liittoa ja Harriet Lonkaa mahdollisuudesta osallistua tähän mielenkiintoiseen hankkeeseen. Kiitän myös kaikkia REKI -työpajoihin osallistuneita mielenkiintoisesta keskustelusta ja hyvistä ideoista. Kiitän myös Geodeettistä laitosta ja heidän työntekijöitään Juha Hyyppää, Sanna Kaasalaista ja Olli Nevalaista. Erityiskiitokset kuuluvat työn ohjaajalleni ja valvojalleni Jussi Leveiselle. Kiitän häntä neuvoista, asiantuntevista kommentteista ja ohjauksesta. Kiitokset myös opiskelija kollegoilleni kannustuksesta ja avusta.*

*Ennen kaikkea suurin kiitos kuuluu perheelleni ja ystävilleni, jotka ovat olleet tukemassa minua koko tämän työn ajan. Haluan myös kiittää Annia lopullisen työn oikolukemisesta ja siskoani Hannaa tiivistelmän tarkastamisesta. Suurimmat kiitokset kuuluvat rakkaalle vaimolleni Sannille, joka on mahdollistanut tasapainottelun opiskelun ja kotielämän välillä.*

*Vielä lopuksi haluan kiittää rakasta poikaani Onnia siitä, että hän on antanut minun nukkua yöni hyvin.*

Espoossa 25.2.2016

Lasse Kangas

# Sisällysluettelo

Tiivistelmä	
Abstract	
Alkusanat	
Sisällysluettelo .....	5
Johdanto .....	7
1.1 Tutkimuksen tausta .....	7
1.2 Tutkimusongelma.....	8
2 Hyperspektrinen lidar-kartoitus .....	10
2.1 Hyperspektrikuvantaminen .....	10
2.1.1 Heijastuminen ja absorptio .....	10
2.1.2 Spektriseen sormenjälkeen vaikuttavat tekijät.....	13
2.1.3 Hyperspektrikuva.....	15
2.1.4 Aktiivinen ja passiivinen hyperspektrikuvantaminen.....	17
2.2 Lidar .....	17
2.2.1 Menetelmä .....	17
2.2.2 Laitteisto .....	19
3 Kiviaineksen laatuun vaikuttavat tekijät .....	21
4 Hyperspektridatan käsittely .....	25
4.1 Datan korjaukset.....	25
4.1.1 Ilmakehäkorjaus.....	25
4.1.2 Jatkumon poistaminen .....	26
4.2 Piirteiden vähentäminen ja muuntaminen.....	27
4.2.1 Sequential Forward Selection .....	28
4.2.2 Sequential Backward Selection .....	28
4.2.3 Branch and bound .....	28
4.2.4 Plus l - take away r.....	29
4.2.5 Pääkomponenttimuunnos.....	29
4.3 Luokittelualgoritmit .....	31
4.3.1 Spectral Unmixing .....	31
4.3.2 Spectral Feature Fitting.....	31
4.3.3 Spektrikulman mukainen luokittelu.....	32
4.3.4 Korrelaatioluokittelu .....	33
4.3.5 Etäisyysluokittelu.....	33
4.3.6 Ryhmittelyanalyysi .....	33
4.3.7 Kanoninen erotteluanalyysi .....	34
5 Näytteet ja menetelmät .....	36
5.1 Näytteet .....	36
5.2 Mittaukset.....	39
5.3 Datan käsittely.....	39
5.4 Harjoitussarjat ja tuntemattomat havainnot .....	40
5.5 Luokittelun tarkkuuden arviointi.....	43
6 Tulokset .....	45
6.1 Spektrikulman mukainen luokittelu .....	45
6.2 Etäisyysluokittelu .....	47
6.3 Korrelaatioluokittelu .....	49
6.4 Kanoninen erotteluanalyysi.....	50

6.5	Ryhmittelyanalyysi .....	52
6.6	Pääkomponenttimuunnos .....	54
7	Tulosten tarkastelu ja pohdinta .....	56
8	Yhteenveto .....	59
	Lähteluettelo .....	60
	Liiteluettelo .....	65
	Liitteet	

## Johdanto

### 1.1 Tutkimuksen tausta

Tämä diplomityö tehtiin osana *Resurssitehokkaan kiviaineshuollon ja pohjavesien suojelun avoin kehittämisympäristö laajalle metropolialueelle (REKI)* -esiselvitystä. Hankkeen tarkoituksena on toteuttaa esiselvitys, joka luo pohjan laajalle ympäristötekniikan konsortiohankkeelle. REKI-hanke on osa Uusimaa-ohjelmassa nimettyä *Kohti Etelä-Suomen vesien ja Itämeren parempaa tilaa (VEDET)* -kärkihanketta. VEDET-hanke käynnistyi vuonna 2013 ja se noudattaa Uusimaa-ohjelman strategisia valintoja *Kasvun mahdollisuudet* ja *Kestävä luonnontalous*.

VEDET-hankkeen yhteyteen on luotu neljä erillistä projektia, jotka liittyvät pohjavesien suojeluun. Koska arvokkaimmat pohjavesialueet sijaitsevat jääkausien synnyttämien harjuja reunamuodostumien yhteydessä, joissa myös helpoiten käytettävät irtomaa-ainekset sijaitsevat, on luontevaa että kiviaineshuolto on osana tarkasteltavaa kokonaisuutta. Uudenmaan alueella hyödynnettävissä olevat soraesiintymät on jo pääsääntöisesti käytetty ja kiviainesten kulutus on siirtynyt enemmissä määrin kalliomurskeisiin.

Suomen sääolosuhteista johtuen rakentamiseen kohdistuu roudankestävyyteen liittyviä vaatimuksia. Sekoittuneet ja hienoainesta sisältävät maa-ainekset ovat tunnetusti herkkiä routimiselle ja siitä johtuen maanvaihtoja joudutaan tekemään sekä talorakentamisen että infrarakentamisen yhteydessä runsaasti. Tästä ja pitkistä välimatkoista johtuen kiviaineksen kulutus on maassamme erittäin runsasta. Suurimmat paineet kohdistuvat tiuhaan asutetulle metropolialueelle, jossa tonttipulan johdosta joudutaan rakentamaan myös heikosti rakentamiseen soveltuvalla maaperällä.

Osana REKI-hanketta on kiviaineksen laatuun liittyvä tarkastelu, jossa on tarkoitus kartoittaa tulevaisuuden tarpeita erilaatuisille kiviaineksille. Koska erityisesti soran saatavuus on jo nyt selvästi heikentynyt ja kuljetusmatkat ovat kasvaneet erittäin pitkiksi, on helpommin saatavilla olevaan kalliokiviainekseen tukeuduttava entistä enemmän tulevaisuudessa. Kalliokiviaines ei kuitenkaan ole yleensä homogeenistä edes yhden louhimon alueella, mikä aiheuttaa laadunhallinnalle haasteita. Laadunvaihtelut saattavat louhimoilla jäädä huomaamatta tai ne huomataan liian myöhään. Syynä tähän ovat osittain kiviaines-standardeissa määritellyt laadunvalvontaan liittyvät tarkastukset, joita tulee tiettyjen laatuparametrien osalta suorittaa erittäin harvoin. Puutteellinen laadunvalvonta saattaa johtaa siihen, että käyttökohteeseen hankittu kiviaines ei toteutakaan sille asetettuja vaatimuksia ja lopputuotteen, kuten tienpinnoitteen elinikä jääkin suunniteltua lyhemmäksi.

Perinteisten laadunhallinnan työkalujen rinnalle on nousemassa automatisoituja laadunvarmistusmenetelmiä, joiden etuna on reaaliaikainen *in-situ* tai *on-site* mittaus. Erilaiset spektroskopiaan tai hyperspektrikuvantamiseen perustuvat menetelmät kykenevät antamaan tietoa mitattavan kohteen atomi-, molekyyli- tai mineraalikoostumuksesta. Käynnissä on myös hankkeita, joissa on tarkoitus selvittää näiden menetelmien soveltuvuutta kiviaineksen laatuparametrien estimoimiseen (Chesner, McMillan 2012). Erääksi vaihtoehdoksi on noussut hyperspektrikuvantaminen, jonka käyttöä mineraalien tunnistuksessa on tutkittu jo 1970-luvulta lähtien (Hunt 1977).

Geodeettinen laitos (nykyisin Paikkatietokeskus FGI) on ensimmäisenä maailmassa kehittänyt hyperspektrisen superjatkumo-lidarin, jonka toiminta perustuu superjatkumolaserin käyttöön valonlähteenä. Hyperspektrilidarin avulla saadaan etäisyystiedon lisäksi kohteesta selville sen heijastusspektritietoa, jonka avulla voidaan tehdä kohteen luokittelua. Hyperspektrikuvantamista on jo aiemminkin käytetty kiviainesten ja mineraalien tunnistamisessa, mutta perinteisesti mittauksissa on valonlähteenä käytetty aurinkoa. Myös niin sanottuun aktiiviseen hyperspektrikuvantamiseen on kehitetty menetelmiä, mutta niiden avulla ei ole ollut mahdollista saada tietoa samanaikaisesti sekä kohteen mineraalikoostumuksesta että rakenteesta.

## 1.2 Tutkimusongelma

Kiviaineshuollossa käytetyn kiviaineksen laadun arvioimiseen käytetyt menetelmät ovat perinteisesti vaatineet geologien katselmuksia ja laboratoriossa standardien mukaisesti suoritettuja tutkimuksia. Käyttökohteen vaativuudesta riippuen testitulosten saaminen kestää yleensä muutamista tunneista muutama viikkoihin. Laboratoriossa tehtäviä analyysejä tehdään yleensä muutamista edustavista näytteistä, joten näytteenottajan ammattitaito ja otettujen näytteiden määrä vaikuttaa olennaisesti saatujen tulosten oikeellisuuteen. Kiviaineksen otonaikainen laadunvarmistus on useampien laatuparametrien osalta hyvin harvoin toistuvaa, joten louhinnan tai oton edistyessä tapahtuvat laatumuutokset saattavat jäädä huomaamatta tai ne huomataan liian myöhään. Tämä saattaa johtaa tilanteeseen, jossa suunniteltu rakenne ei olekaan niin pitkäikäinen kuin oli suunniteltu. (Chesner, McMillan 2012)

Toinen laadun määrittelyyn liittyvä ongelma on liian hyvälaatuisen kiviaineksen käyttö toissijaisissa kohteissa. Kiviaineksen epätasainen alueellinen jakautuminen on johtanut siihen, että kaikilla alueilla ei ole käytössä korkealuokkaista kovakiveä tai sen varannot ovat huvenneet. Jos kiviaineslouhoksen kiviaineksen laatua ei ole määritetty riittävän hyvin, saattaa olla että korkealuokkaista kovakiveä myydään täyttö- tai pohjarakennusmateriaaliksi kohteeseen, jossa heikkolaatuisempikin aines olisi riittänyt. Näin ollen tuhlaataan hyvälaatuisia materiaalia toissijaiseen kohteeseen.

Kiviaineshuollon tehostamisen kannalta nopeat ja usein toistuvat laadunhallinnan menetelmät ovat tulevaisuudessa ehdottoman tärkeitä. Näytteiden reaaliaikainen tutkiminen paikallaan toisi joustavuutta, nopeutta ja tehokkuutta kiviaineshuoltoon. Nykyiset alkuaine-, molekyyli- ja hyperspektrianalyysiin soveltuvat tekniikat ovat kehittyneet sille asteelle, että kenttäolosuhteisiin sopivia laitteistoja on mahdollista toteuttaa.

Mittaustulosten suora tulkinta on varsinkin massamittauksissa, kuten laadunvarmistuksessa, liian hidasta ja työlästä. Tästä johtuen tulosten tulkinnassa ja analysoinnissa käytetään fyysikaalisiin ominaisuuksiin ja tilastotieteisiin pohjautuvia menetelmiä. Luokittelualgoritmit perustuvat kriteerifunktioiden käyttöön ja niiden tarkoituksena on löytää mahdollisimman hyvin referenssispektrejä vastaavia pikseleitä. Myös niin sanottuja *spectral unmixing* -tekniikoita käytetään erottelemaan sekoittuneen spektrin päätejäseniä. Tämän työn analyyseissa käytetyt luokittelualgoritmit ovat spektrikulman mukainen luokittelu (SAM), korrelaatioluokittelu (SCM), etäisyysluokittelu, ryhmittelyanalyysi ja kanoninen erotteluanalyysi. Lisäksi dimensioiden vähentämisen vaikutuksia tarkastellaan pääkomponenttimuunneltujen muuttujien ja ryhmittelyanalyysin avulla.



Tämän työn ensisijaisena tarkoituksena on selvittää, onko hyperspektrilidarin ja analyysimenetelmien avulla mahdollista luokitella kiviaineita ennalta määrättyihin luokkiin. Työn käytännön osuus toteutettiin Geodeettisen laitoksen kehittämällä hyperspektrisellä superjatkumo-lidarilla. Tutkittavina kivilä oli kaupallisia, yleisessä käytössä olevia rakennuskiviä. Tutkittavat kivet olivat migmatiitti, gneissi, kvartsiitti ja neljä erilaista graniittia.

Työn teoriaosuudessa käsitellään myös kiviaineksen laatuun liittyviä tekijöitä ja niiden yhteyttä spektriseen sormenjälkeen. Tämän tarkoituksena on osoittaa, että ainakin teoriassa kiviaineksen laatuparametreja, kuten lujuusominaisuuksia, on mahdollista johtaa spektritiedosta. Kaikkia tässä työssä esiteltyä datan käsittely- ja analysointimenetelmiä ei käytetty testidatalle. Tämä johtuu testidatan mittausten menetelmästä ja laitteiston rajoituksista. Niiden esittelyä voi kuitenkin pitää perusteltuina, koska mahdollisissa jatkotutkimuksissa muun muassa piirteidenvalinta saattaa olla merkittävässä roolissa. Tässä työssä ei tulla keskittymään lidarin antamaan paikkatietoon muuten kuin perusteorian osalta.

Työn toisessa kappaleessa käsitellään teoriaa hyperspektrikuvantamisen, lidar-kartoitusten ja mittauksissa käytetyn laitteiston osalta. Kolmannessa kappaleessa käsitellään kiviaineksen laatuun vaikuttavia tekijöitä ja niiden yhteyttä heijastusspektriin. Työn neljännessä kappaleessa keskitytään spektritiedon analysointimenetelmiin ja teoriaan niiden taustalla. Viidennessä kappaleessa tutustutaan tehtyihin mittauksiin, tutkimusaineistoon ja sen esikäsittelyyn, sekä harjoitussarjojen ja tuntemattomien havaintojen valintaan. Myös luokittelutarkkuuden arvioinnissa käytettävät menetelmät esitellään viidennessä kappaleessa. Kuudes kappale sisältää analyysien tulokset. Seitsemännessä kappaleessa keskitytään saatujen tulosten tarkasteluun ja pohdintaan. Työn yhteenveto löytyy kahdeksannesta kappaleesta.

## 2 Hyperspektrinen lidar-kartoitus

Hyperspektrikuvantaminen on spektroskopian osa-alue, jossa sähkömagneettisen säteilyn vuorovaikutus tutkittavan kohteen kanssa aiheuttaa heijastusspektrissä havaittavia absorptiopiirteitä. Tutkittavassa kohteessa tapahtuvien atomiprosessien vaihtelevuus ja aallonpituusriippuvuus ovat keskiössä informaation johtamisessa heijastusspektristä. Hyperspektrinen lidar yhdistää hyperspektrikuvantamisen tuottaman spektritiedon lidarin tuottamaan spatioaliseen tietoon ja tarjoaa näin ollen monipuolisen välineen kaukokartoitukseen.

Hyperspektrikuvantaminen ja lidar ovat itsenäisinä tekniikoina olleet jo kauan käytössä. Hyperspektrikuvantamiselle ja niin kutsutulle *spektriselle sormenjäljelle* pohjaa loivat Graham R. Huntin ja John W. Salisburyn työryhmä, joka julkaisi useita mineraalien ja kivien heijastusspektriin liittyviä artikkeleita 1970-luvulla. He loivat teorian, jonka mukaan kvanttimekaaniset prosessit, kuten elektroni- ja värähtelyprosessit, määrittävät havaitun heijastusspektrin. (Hunt 1977)

Lidar-laitteistot kehitettiin vähän laservalon keksimisen jälkeen 1960-luvulla. Ensimmäiset käyttökohteet olivat meteorologiassa, jossa lidaria käytettiin pilvien mittaamiseen. (Reagan, McCormick et al. 1989) Lidarin käyttö alkoi selkeästi lisääntyä 1990-luvun lopulla, kun markkinoille saatiin vastaanotinsensoreita, jotka pystyvät käsittelemään tuhansia pulsseja sekunnissa. GIS-ohjelmistojen kehittyminen, laskentatehon parantuminen ja muistikapasiteetin lisääntyminen 2000-luvulla mahdollistivat lidar-mittausten laaja-alaisen käytön.

Seuraavaksi tässä työssä esitellään hyperspektrikuvantaminen ja lidar itsenäisinä tekniikoina. Hyperspektrikuvantamisen osalta käydään läpi spektrinmuodostumisen taustalla olevaa fysiikkaa ja erityinen painoarvo annetaan kivien ja mineraalien spektriseen tunnistamiseen vaikuttaville tekijöille. Lidarin osalta käsitellään perusteet lidar-mittauksista sekä tässä tutkimuksessa käytetystä superjatkumovalonlähteestä.

### 2.1 Hyperspektrikuvantaminen

Hyperspektrikuvantamisella tarkoitetaan informaation keräämistä ja prosessointia näkyvää valoa laajemmalla sähkömagneettisen spektrin alueella. Perinteinen valokuvaus perustuu kolmen aallonpituuskaistan (punainen, vihreä ja sininen) tallentamiseen ja käsittelemiseen. Hyperspektri- ja multispektrikuvantamisessa spektritietoa kerätään useammalta eri aallonpituudelta ja leveämmältä aallonpituusalueelta. Mineraalien tutkimisessa käytetyt aallonpituudet asettuvat yleensä näkyvän valon ja lyhytaaltoisen infrapun välille (325–2500 nm) (Hunt 1977). Multispektrikuvantamisen erottaa hyperspektrikuvantamisesta kanavien lukumäärä ja diskreettisyys, eli multispektrikuvantamisessa ei tuoteta kohteelle jatkuvaa spektriä.

#### 2.1.1 Heijastuminen ja absorptio

Kun valo kohtaa esteen, se joko heijastuu, absorboituu tai läpäisee sen. Energian säilymislain perusteella intensiteetille (I) saadaan:

$$I(\lambda) = R(\lambda) + A(\lambda) + T(\lambda) = 1 \Rightarrow R(\lambda) = 1 - A(\lambda) - T(\lambda), \quad (1)$$

missä  $R$  on reflektanssi,  $A$  on absorptio ja  $T$  on läpäisykyky. Kuten kaavasta 1 voidaan huomata, valon käyttäytyminen riippuu kohteen ominaisuuksista ja siihen osuvan valon aallonpituudesta. Kokonaisintensiteetin komponenttien suhde siis määrittelee havaitun heijastuspektrin kyseisellä aallonpituudella. Reflektanssi  $R$  on määritelty kohteeseen saapuvan ja siitä takaisin heijastuvan säteilyn osamääränä (van der Meer 2004):

$$R = \frac{I_r(\lambda)}{I_i(\lambda)}, \quad (2)$$

missä  $I_r$  on heijastuneen säteilyn intensiteetti ja  $I_i$  on kohteeseen saapuneen säteilyn intensiteetti. Absorptio määritellään vastaavanlaisesti:

$$A = \frac{I_a(\lambda)}{I_i(\lambda)}, \quad (3)$$

missä  $I_a$  on absorboituneen säteilyn intensiteetti ja läpäisykyky:

$$T = \frac{I_t(\lambda)}{I_i(\lambda)}, \quad (4)$$

missä  $I_t$  on läpäisseen säteilyn intensiteetti.

Mineraalien heijastusominaisuudet riippuvat niiden kemiallisesta koostumuksesta ja kiderakenteesta. Useille mineraaleille on löydetty niin sanottu spektrinen sormenjälki, jonka avulla niiden tunnistaminen on mahdollista. Selkeimmät absorptiopiirteet näkyvän valon ja infrapunasäteilyn alueella (~400–1000 nm) johtuvat siirtymämetallien, kuten raudan, kromin tai nikkelin aiheuttamista leveistä absorptiokaistoista. Pidemmällä aallonpituuksilla (~1000–3000 nm) absorptiopiirteitä dominoivat veden, karbonaattien, sulfaattien tai hydroksidien värähtelystä ja rotaatiosta johtuvat jyrkät absorptiokaistat. (Hunt 1977, Leverington 2009, van der Meer 2004)

Kohdetta valaistaessa tapahtuu kahdenlaista heijastumista (Kuva 1). Diffuusi heijastus tapahtuu karkeilla pinnoilla, joiden korkeusvaihtelut ja osasten koko ovat suurempia kuin säteilyn aallonpituus. Diffuusissa heijastuksessa säteily heijastuu kaikkiin suuntiin. Diffuusin heijastuksen intensiteetti noudattaa käytännössä Lambertin kosinilakia. Lain mukaan heijastuneen säteilyn intensiteetti on suoraan verrannollinen tulokulman ja pinnan normaalin välisen kulman kosiniin. Koska intensiteetti on myös kääntäen verrannollinen kohteen etäisyyden neliöön, diffuusin heijastumisen intensiteetiksi ( $I_d$ ) saadaan:

$$I_d = \frac{R_d \times \cos \theta}{x^2}, \quad (5)$$

missä  $R_d$  on diffuusi heijastuskerroin,  $\theta$  on säteen tulokulman ja pinnan normaalin välinen kulma ja  $x$  on kohteen ja tarkkailupisteen välinen etäisyys. Kaavasta 5 voidaan helposti huomata, että diffuusin heijastuksen intensiteetti on suurimmilla, kun tuleva säde on heijastavan pinnan normaalin suuntainen. (Hasegawa 2006)

Peiliheijastus tapahtuu ilman sirontaa tai diffuusiota. Peiliheijastuksessa lähes kaikki tuleva säteily heijastuu pois samansuuruisessa kulmassa pinnan normaaliin nähden kuin tulokulma ja sen intensiteetti laskee nopeasti kulman muuttuessa. Peiliheijastus on tulokulman

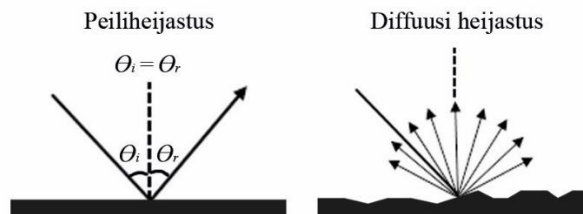
ja materiaalin ominaisuuksien välinen funktio, ja sen intensiteetti ( $I_s$ ) voidaan lausua seuraavasti:

$$I_s = \frac{W(\theta) \times \cos^n(\gamma)}{x^2}, \quad (6)$$

missä  $W(\theta)$  on peiliheijastuskerroin tulokulman ( $\theta$ ) funktiona,  $\gamma$  on heijastuskulman ja tarkkailukulman välinen kulma,  $n$  määrittää peiliheijastuksen terävyyden (yleensä välillä 1-200) ja  $x$  on kohteen ja tarkkailupisteen välinen etäisyys. (Hasegawa 2006)

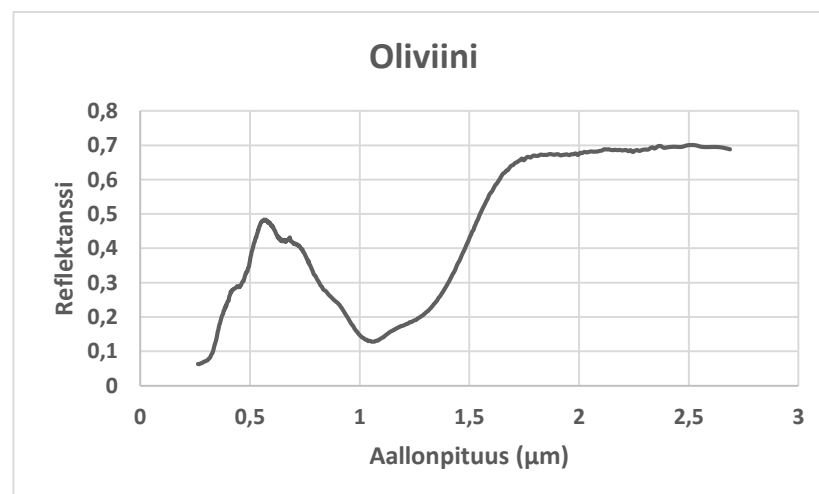
Lidar-mittauksia varten yhtälö voidaan yleensä yksinkertaistaa olettamalla, että valonlähde ja tarkkailupiste ovat samat ja olettamalla peiliheijastuskerroin vakioksi. Näin ollen peiliheijastuksen intensiteetille saadaan (Hasegawa 2006):

$$I_s = \frac{W \times \cos^n(2\theta)}{x^2}, \quad (7)$$



Kuva 1. Peili- ja diffuusiheijastus.

Absorptio on kvanttimekaaninen prosessi, jossa foton siirtää energian atomille, jonka valenssielektronit virittyvät korkeampaan energiatilaan. Korkeampi energiatila ei ole vakaa, joten energiatila purkautuu yleensä lämpöenergiana tai absorboitunut foton saattaa emittoitua takaisin fotonina. Atomi voi absorboida vain fotonin, jonka energia vastaa atomin nykyisen energiatilan ja jonkin sille ominaisen korkeamman energiatilan erotusta. Tästä johtuvat heijastusspektrissä näkyvät kullekin aineelle ominaiset absorptiokaistat. Kuvassa 2 nähdään oliviinin heijastusspektri, jossa aallonpituuksilla 700 nm-1400 nm spektriä hallitsee  $\text{Fe}^{2+}$ -ionien aiheuttama leveä absorptiokaista (Hunt 1977, Meer, Jong 2002).



Kuva 2. Oliiviinin heijastusspektri (USGS, GDS70.b Fo89 115μm).

Absorptiokaistan syvyys korreloi läsnä olevan materiaalin suhteellisen määrän kanssa. Tämä korostuu erityisesti monomineraalien kohdalla. Myös raekoko vaikuttaa absorptiokaistan syvyyteen. (Guha, Rao et al. 2012)

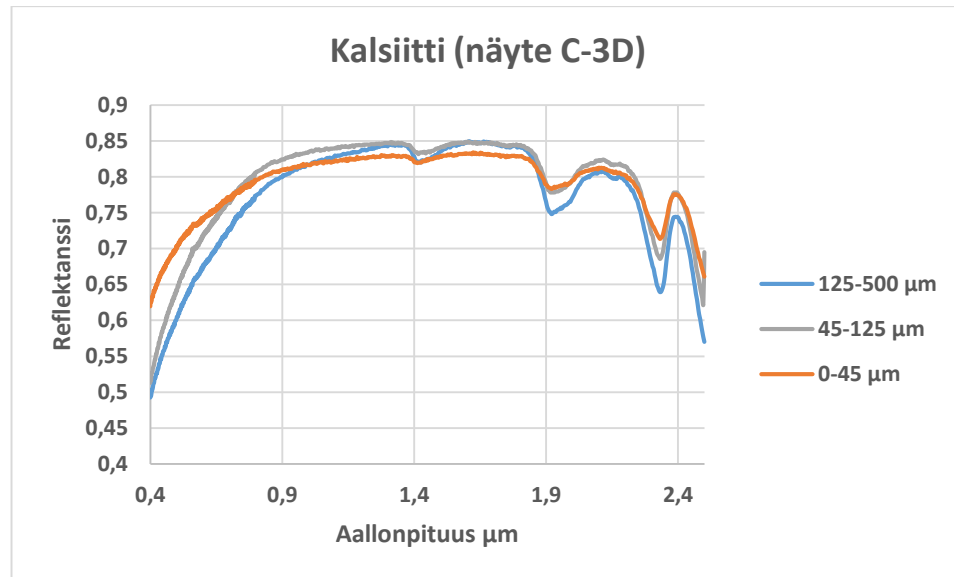
Absorptioprosessit voidaan jakaa elektronisiin sekä rotaatio- ja värähtelyprosesseihin. Koska elektroniset prosessit vaativat suuremman energiamäärän toteutuakseen ( $E = hc/\lambda$ , missä  $h$  ja  $c$  ovat vakioita), tapahtuvat ne lyhemmillä aallonpituuksilla. Tyypillisimmät elektroniset prosessit johtuvat elektronien virittymisestä korkeammalle energiatasolle. Ne voivat myös johtua varausten siirtymisestä tai alkuaineiden välisestä siirroksista, jossa fotoni saa elektronin siirtymään ionien tai ionien ja ligandien välillä. Koska atomien ja ionien energiatilat ovat diskreettejä, on niiden virittymiseen liittyvä aallonpituusriippuvuus myös diskreetti. (Hunt 1977)

Rotaatio- ja värähtelyprosessit johtuvat molekyylien tai kiderakenteen sidosten rotaatiosta ja värähtelystä. Sidoksia voidaan kuvailla jousina joiden päässä on painot. Riippuen sidosten voimakkuudesta ja niiden massasta, erilaiset molekyylit värähtelevät eri taajuuksilla. (Meer, Jong 2002) Värähtelyn aiheuttamat absorptiokaistat ovat kapeita (10-20nm) verrattuna elektronisten prosessien tuottamiin absorptiokaistoihin (Hunt 1977).

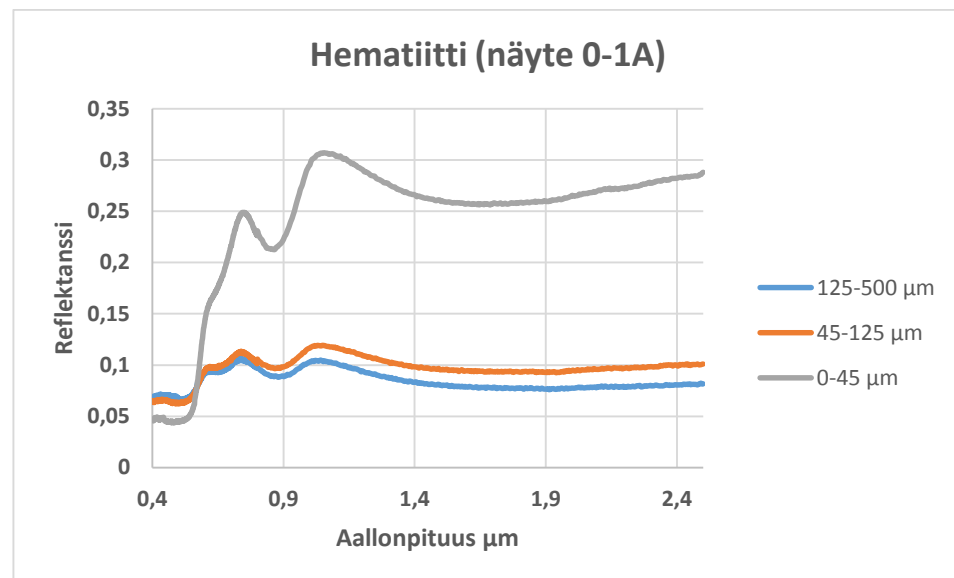
### 2.1.2 Spektriseen sormenjälkeen vaikuttavat tekijät

Spektriseen sormenjälkeen vaikuttaa alkuaineiden ja molekyylien aiheuttamien elektronisten ja värähtelyprosessien lisäksi myös kiderakenteen hilavirheet ja raekoko sekä mittaus- tekniset tekijät kuten vaihekulma ja näytteenoton pinta-ala (säteen koko). Kivien ja mineraalien heijastusspektrien käyttäytyminen raekoon tai mittaus-tekniikoiden tekijöiden vaikutuksesta ovat hyvin samankaltaisia (Guha, Rao et al. 2012). Seuraavaksi käsitellään yleisimmät spektriseen sormenjälkeen vaikuttavat tekijät ja osoitetaan, että kivienlajien tunnistamisessa on parempi käyttää kiviläytteistä johdettuja referenssispektrejä kuin niiden päämineraaleista johdettuja kompositiospektrejä. Ilmakehän vaikutuksia heijastusspektriin käsitellään kappaleessa 4.1.1. Tässä kappaleessa olevat kuvat on tehty käyttämällä Aster 2.0 – spektrikirjastosta saatuja dataa (Baldrige, Hook et al. 2009).

Hyvin kiteytyneet mineraalit eroavat heikosti kiteytyneistä mineraaleista pääsääntöisesti korkeamman taustaheijastavuuden omaavina. Mineraalien ja kivien raekoko ja asu vaikuttavat heijastusspektrin intensiteettiin ja absorptiokaistojen jyrkkyyksiin, mutta eivät siirrä kaistojen sijaintia (Guha, Rao et al. 2012). Yleisesti ottaen karkearakeiset näytteet heijastavat heikommin, johtuen sähkömagneettisen säteilyn kulkeman matkan pituudesta (Beerin laki) ja pintaheijastuksen määrästä (van der Meer 2004). Pienillä rakeilla on enemmän pinta-alaa suhteessa tilavuuteen, joten niiden pinnoilta tapahtuu enemmän sirontaa (Clark 1999). Nyrkkisääntönä voidaan pitää, että raekoon pienentyessä heijastuksen intensiteetti kasvaa ja absorptiokaistojen syvyys pienenee (Cloutis 1996). Kuvissa 3 ja 4 on nähtävissä kalsiitin ja hematiitin kolmeen eri karkeuteen jauhettujen näytteiden heijastusspektrit, joista on huomattavissa eroja reflektanssissa. Merkillepantavaa on, että korkean reflektanssin omaava kalsiitti käyttäytyy raekoon muuttuessa huomattavasti hienovaraisemmin kuin hematiitti. Molempien näytteiden kohdalla voidaan huomata, ettei reflektanssin muutos ole täysin lineaarista suhteessa raekokoon.



Kuva 3. Kalsiitin heijastusspektri kolmeen eri kärkeuteen jauhetuista näytteistä (Baldridge, Hook et al. 2009).

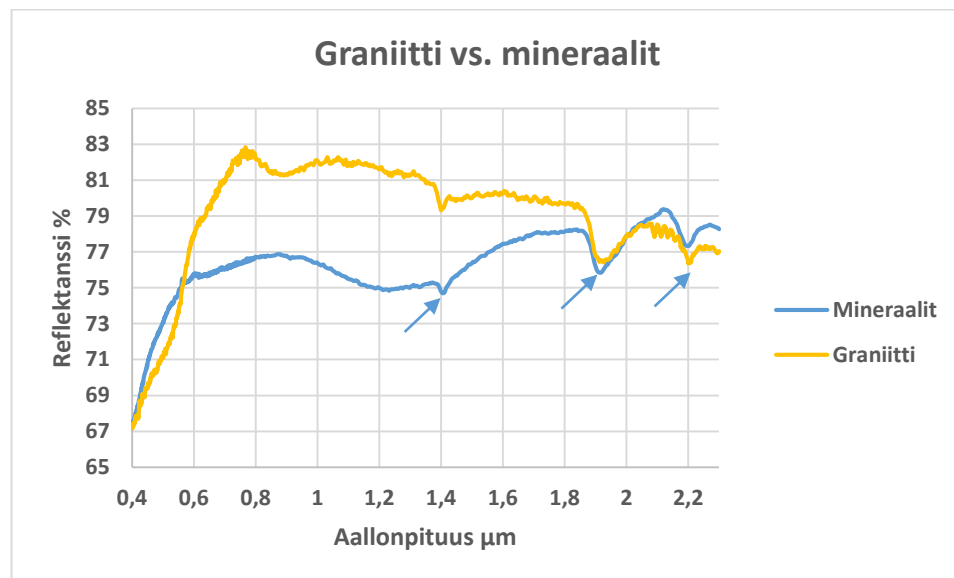


Kuva 4. Hematiitin heijastusspektri kolmeen eri kärkeuteen jauhetuista näytteistä (Baldridge, Hook et al. 2009).

Mittausalueen pinta-ala, joka on yleensä suhteessa mittausetäisyyteen, vaikuttaa heijastusspektriin hyvin samankaltaisesti kuin raekoko. Jos kivilaatu tai mineraalikoostumus ei muutu olennaisesti mitattavan pinta-alan kasvaessa, säilyvät spektrin ominaispiirteet selvästi erotettavina. Mittausten vaihekulmalla, eli valonlähteen ja vastaanotinoptiikan välisen kulman muutoksilla on myös vaikutusta havaittuun spektriin. Vaihekulman muutokset vaikuttavat taustaheijastavuuteen ja spektrin profiiliin, mutta eivät siirrä absorptiokaistojen paikkoja. (Guha, Rao et al. 2012)

Kivien heijastusspektriä hallitsevat sen päämineraalien absorptiopiirteet. Kivien heijastusspektri ei kuitenkaan ole suora lineaarinen kombinaatio sen mineraalikoostumuksesta, vaan absorptiokaistojen paikat ja syvyydet muokkautuvat epälineaarisen spektrin sekoittumisen

kautta.(Guha, Rao et al. 2012, Cloutis 1996) Kuvassa 5 on nähtävissä graniitin ja sen päämineraalien yhdistelmän (lineaarinen sekoitus) spektrit, jossa graniitin heijastavuus on skaalattu suuremmaksi helpomman havainnollistamisen vuoksi. Kuvasta 5 on havaittavissa, että absorptiokaistojen paikat ja muodot toistuvat kohtuullisen hyvin, mutta eivät täydellisesti. Absorptiokaistojen minimiit ovat graniitin heijastusspektrissä noin 4nm lyhemmillä aallonpituuksilla. Myös absorptiokaistojen syvyyksissä ja symmetrioissa on havaittavissa eroja. Tämä johtuu siitä, että kivimatriisissa hyvin lähekkäin olevat mineraalirakeet vaikuttavat toistensa absorptio-ominaisuuksiin (Meer, Jong 2002).

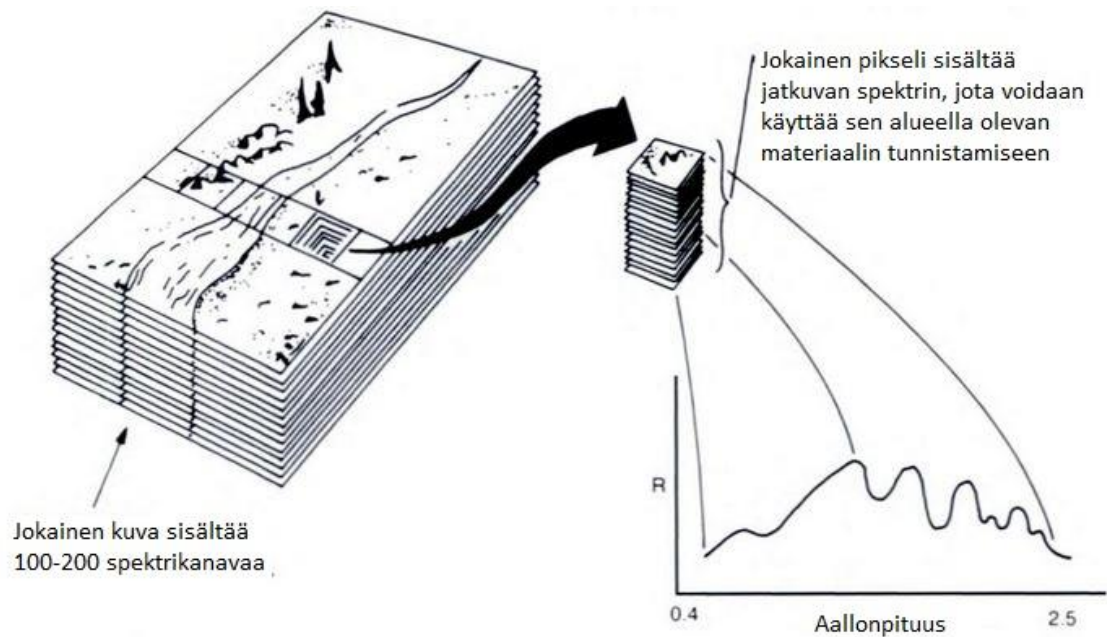


Kuva 5. Graniitin heijastusspektri ja sen päämineraaleista koostettu kompositiospektri (Baldrige, Hook et al. 2009).

Koska kivien spektriset sormenjäljet sisältävät selkeitä ominaispiirteitä, eivätkä ne ole mineraaleja herkempiä fysikaalisten tai mittausteknisten muutosten aiheuttamille häiriöille, on Guha et al. (2012) mukaan parempi käyttää referenssispektreinä kiviaineksista mitattuja spektrejä, kuin päämineraaleista johdettuja kompositiospektrejä. Referenssispektrejä käytettäessä on kuitenkin huomattava, että laboratorioissa tai kentällä tuotetut referenssispektrit ovat aina riippuvaisia olosuhteista ja mittauslaitteistoista, joten niitä voidaan käyttää ainoastaan approksimaatioina kyseisen aineen spektristä (Schowengerdt 2007). Edes samasta materiaalista laboratorioissa mitatut spektrit eivät yleensä ole täysin identtisiä, johtuen näytteiden pinnanmuotojen vaihtelevuudesta (Manolakis, Shaw 2002).

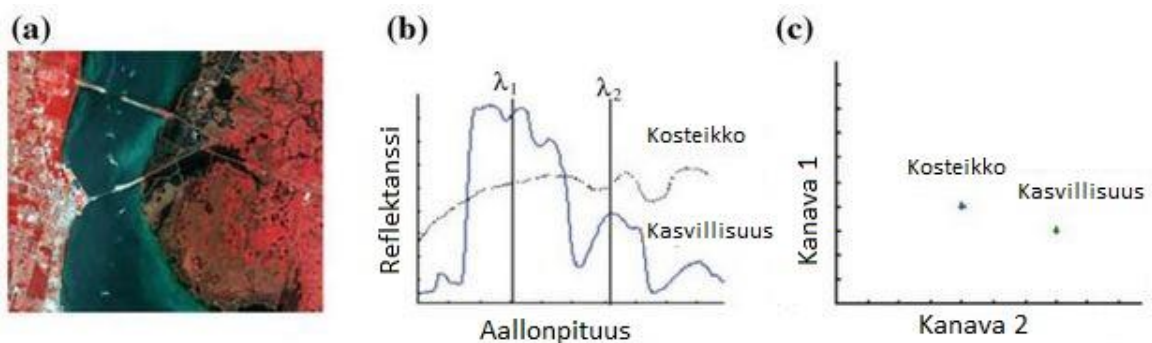
### 2.1.3 Hyperspektrikuva

Hyperspektrikuva on useista toistensa läheisistä kanavista (aallonpituuksista) koottu kuva, joka sisältää informaatiota näkyvää valoa laajemmalla sähkömagneettisen säteilyn alueella.(Manolakis, Shaw 2002) Tietokoneiden ja näyttölaitteiden käyttämät RGB – kuvat sisältävät vain kolme kanavaa (punainen, vihreä ja sininen), joiden aallonpituudet asettuvat samalle alueelle ihmisen näkökyvyn kanssa (n. 400–700 nm). Jokainen näistä kanavista saa numeroarvon, jotka yhdessä määrittelevät toistettavan pikselin värin. Kuva taas rakentuu pikseleistä, joista jokaisella on oma paikkakoordinaattinsa.



Kuva 6. Hyperspektrikuvaan koostumus (Barrett 2013).

Toisin kuin tavallinen RGB-kuva, hyperspektrikuva (kuva 6) sisältää useita (kymmenistä satoihin) erittäin kapeita, toisiaan lähellä olevia kanavia, joiden aallonpituusulottuvuus on yleensä näkyvää valoa huomattavasti laajemmalla alueella. Hyperspektrikuva voi siis ajatella sisältävän joukon päällekkäisiä harmaasävykuvia (yksikanavaisia), joista jokainen sisältää yhden kapean aallonpituuskanavan sisältämän informaation. Pikselitasolla jokainen hyperspektrikuvaan pikseli sisältää niin monta arvoa, kun kuvassa on kanavia. Yhden pikselin voi siis kuvata vektorina, jonka alkioita nämä arvot ovat (Theodoridis, Koutroumbas 2009). Hyperspektrikuva voidaan visualisoida esimerkiksi kuvan 7 kaltaisena väärävärinä, spektrinä tai muuttujatasolla.



Kuva 7. Väärävärikuva (a), Spektri (b), Muuttujataso (3). (Wang, Zhao 2015)

Luonnollisesti hyperspektrikuvaan sisältämä informaation määrä on selkeästi suurempi kuin tavallisen RGB-kuvan. Useasti vierekkäiset kanavat korreloivat vahvasti keskenään, joten hyperspektrikuvaan dimensioiden vähentäminen on monissa käytännön sovelluksissa järkevää. Dimensioiden vähentämisessä tulee kuitenkin olla varovainen, jos tutkimuskohteelle olennaisista aallonpituusalueista ei olla täysin varmoja. Dimensioiden vähentämisellä saatetaan hukata tärkeää tietoa, jota voitaisiin hyödyntää kohteen luokittelemisessa. Dimensioiden vähentämistä ja piirteiden valintaa käsitellään tarkemmin kappaleessa 4.2.



Hyperspektrikuvia vähemmän, mutta RGB-kuvia enemmän kanavia sisältäviä kuvia kutsutaan multispektrikuviksi. Multispektrikuvat eroavat hyperspektrikuvista vähäisemmän kanavien lukumäärän lisäksi siinä, että niiden kanavat ovat yleensä leveämpiä ja diskreettejä. Monet kaukokartoituslaitteet hyödyntävät multispektrikuvantamista (Camps-Valls 2012).

#### 2.1.4 Aktiivinen ja passiivinen hyperspektrikuvantaminen

Hyperspektrikuvantaminen on perinteisesti suoritettu passiivisesti, eli auringosta tulevaa säteilyä hyväksi käyttäen. Passiivisen kuvantamisen haittoina ovat kuitenkin kolmiulotteisissa kohteissa tapahtuva varjostuminen ja säteilyn voimakkuuden muutokset, jotka johtuvat esimerkiksi säätilasta. Osittain varjostuneiden alueiden tunnistus on käytännössä mahdotonta, koska ne eroavat valaistuista alueista sekä amplitudin että spektrin muodon osalta. Myöskään öisin tapahtuva mittaaminen ei ole ollut mahdollista passiivisilla tekniikoilla. (Nischan, Newbury et al. 2000)

Hyperspektrikuvantamista on tehty myös aktiivisilla menetelmillä, joiden avulla on saatu eliminointia edellä mainittuja heikkouksia. Asettamalla valonlähde ja vastaanotinoptiikka samaan pisteeseen saadaan eliminointia varjostuminen ja pienennettyä mittauskulman vaikutuksista johtuvia eroja eri kuvien välillä. Aktiivisen kuvantamisen etuina on myös se, että käytetyn valonlähteen ominaisuudet tunnetaan hyvin. Aktiiviset kuvausmenetelmät mahdollistavat myös pimeässä suoritettut mittaukset. (Nischan, Newbury et al. 2000)

Valokuitukaapeleiden ja pienikokoisten pulssilasereiden viimeaikainen kehitys on tuonut markkinoille valonlähteitä, jotka ovat yhtä hyvin suuntautuneita kuin laserit ja toimivat erittäin laajan spektrin alueella. Nämä valonlähteet ovat mahdollistaneet aktiivisen hyperspektrikuvantamisen ja lidar-mittausten yhdistämisen. Tässä tutkimuksessa käytetty hyperspekt-rinen superjatkumo-lidar hyödyntää aktiivista kuvantamista.

## 2.2 Lidar

Lidar (engl. Light Detecting and Ranging) on kaukokartoituslaite, joka toimii tutkan (engl. Radio Detecting And Ranging) tavoin, lähettämällä sähkömagneettista säteilyä kohteeseen ja mittaamalla heijastuneen pulssin kulkuajan. Lidar eroaa tutkasta siinä, että sen lähettämä sähkömagneettinen säteily (laservalo) on aallonpituudeltaan paljon lyhyempää kuin tutkan käyttämä radio- tai mikroaaltosäteily. Tästä johtuen lidarilla voidaan mitata huomattavasti pienempiä kohteita kuin tutkalla. (Olsen 2007) Lidar tuottaa 3D-pistepilven, joka yksityiskohtaisesti kuvaa mitattavan kohteen geometriaa. (Alho, Hyyppä et al. 2011). Lidaria käytetään muun muassa maanmittauksessa, geomatiikassa, geologiassa, seismologiassa, metsäntutkimuksessa, meteorologiassa ja arkeologiassa. Seuraavaksi tässä työssä esitellään lidar-mittauksen perusteet ja tässä työssä käytetty laitteisto.

### 2.2.1 Menetelmä

Lidar-mittauksilla saatu pistepilvi koostuu mittauspisteistä, joiden koordinaatit (x,y,z) tunnetaan. Koordinaatit lasketaan kulma- ja etäisyysmittausten perusteella. Etäisyyden määrittäminen perustuu yleensä joko kulkuaikaan tai vaihe-eroon. Kulkuaikaan perustuvissa laitteistoissa mitataan lähetetyn ja palaavan laserpulsin välinen aikaero. Koska valonnopeus tunnetaan, voidaan etäisyys laskea yksinkertaisesti kaavalla:

$$r = \frac{dt \cdot c}{2}, \quad (8)$$

missä  $r$  on etäisyys,  $dt$  on kulunut aika ja  $c$  on valonnopeus. Kuluneen ajan ja valonnopeuden tulo jaetaan kahdella, koska mitattu aika on kulunut pulssin kulkiessa kohteeseen ja sieltä takaisin. (Alho, Hyyppä et al. 2011)

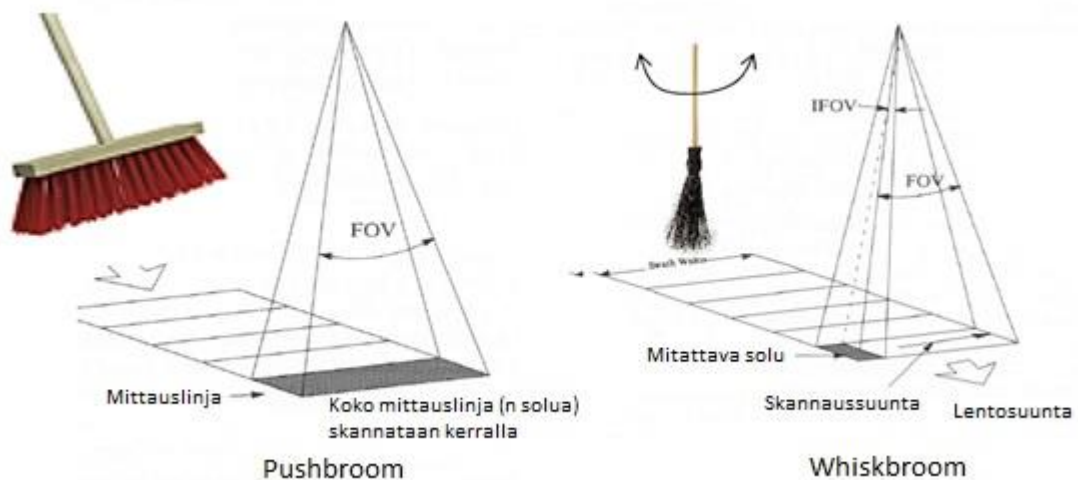
Vaihe-eroon perustuvissa laitteissa kulkuaika lasketaan lähetetyn ja vastaanotetun pulssin vaihe-eron perusteella. Etäisyys voidaan laskea kaavalla:

$$r = \left( \frac{d\phi}{2\pi} \right) \lambda_0, \quad (9)$$

missä  $d\phi$  on vaihe-ero ja  $\lambda_0$  on lyhin moduloitu aallonpituus. Vaihe-eromenetelmällä saadaan mitattua tiheämpää aineistoa kuin kulkuaikaan perustuvalla menetelmällä, koska mittaus on jatkuvaa, eikä palaavaa pulssia tarvitse odottaa ennen uuden lähettämistä. (Alho, Hyyppä et al. 2011) Tässä työssä käytetty laitteisto on kulkuaikaan perustuva.

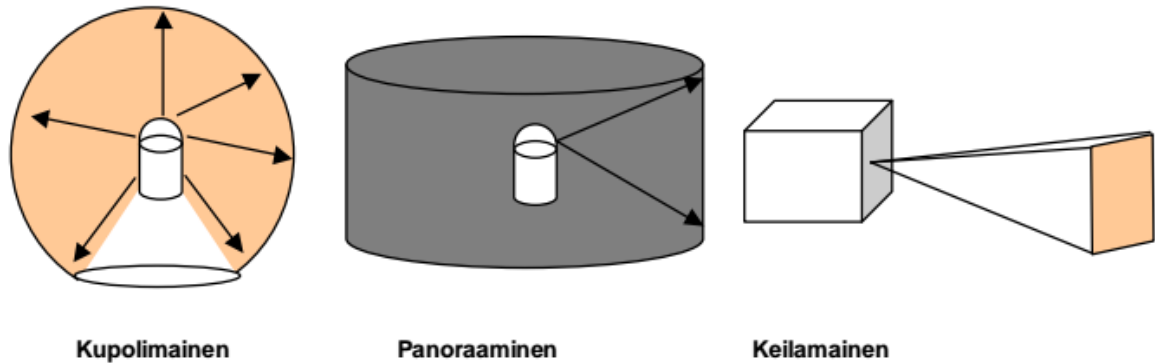
Lidar-mittaukset voidaan jakaa ilmasta (engl. airborne) ja maasta (engl. terrestrial) tehtäviin mittauksiin. Ilmasta tehtävät mittaukset suoritetaan lentokoneesta, helikopterista tai satelliitista. Lidarissa on perinteisesti käytetty valonlähteenä monokromaattista laseria, jonka aallonpituus on käyttötarkoituksesta riippuen yleensä välillä 400–1600 nm. Yleisimmin käytetyt laserit ovat joko 1064 nm diodipumpattu YAG-laseri (Yttrium-Alumiini-Granaatti) tai 532 nm kaksoisdiodipumpattu YAG-laseri. Lyhemmän aallonpituuden (532 nm) laseria on tyypillisesti hyödynnetty batymetrisissä mittauksissa, koska sen vedenläpäisykyky on huomattavasti pitkän aallonpituuden omaavia valonlähteitä parempi. (Vosselman, Maas 2011)

Mittaukset suoritetaan tyypillisesti lentäen mittauslinjoja pitkin, käyttäen joko kohtisuoraan lentosuuntaan nähden toimivaa keilainta (whiskbroom) tai lentosuunnan keilausta (pushbroom) (kuva 8). Lentokoneen paikkatieto (x,y,z) saadaan käyttämällä GPS-järjestelmää ja lentokoneen asennon muutokset mitataan inertian mittausyksikön avulla (IMU engl. Inertia Measurement Unit).



Kuva 8. Pushbroom ja whiskbroom -tekniikat (János, Tünde 2008).

Maasta tehtävissä mittauksissa lidar-laitteisto (maalaserkeilain) on yleensä paikallaan kolmijalan päällä. (Alho, Hyyppä et al. 2011) Mittausgeometria voi olla kupolimainen, panoraaminen tai keilamainen (kuva 9). Kupolimaisessa mittauksessa saadaan mitattua koko ympäröivä alue, lukuun ottamatta pientä katvealuetta laitteen alla. Panoraamisissa mittauksissa laitteen yläpuoliset ja alapuoliset alueet jäävät katveeseen. Keilamaisissa mittauksissa tuotetaan kameramaisia pistepilviä.(Vosselman, Maas 2011)

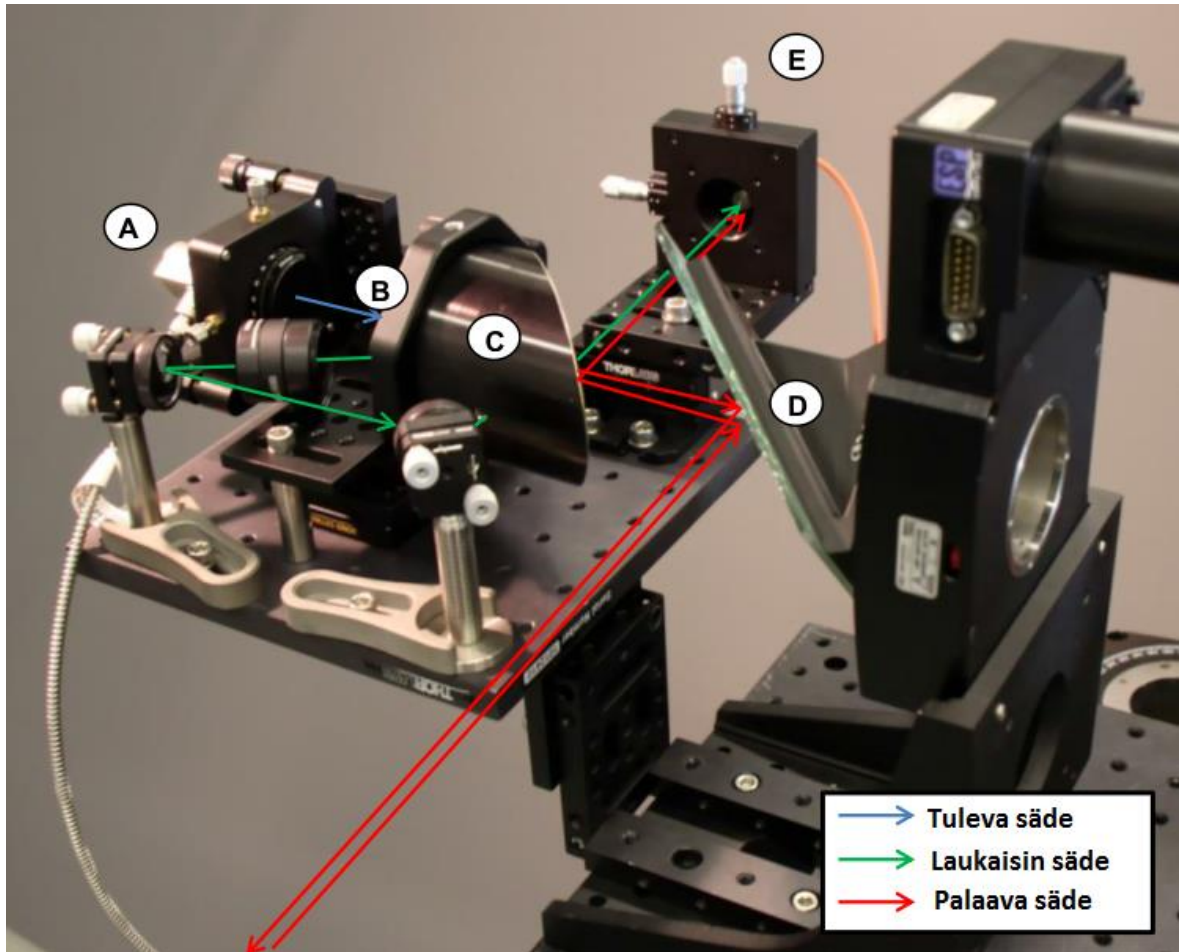


Kuva 9. Mittausgeometriat (Joala 2006).

## 2.2.2 Laitteisto

Lidar-laitteisto koostuu laserista, skannerista, optiikasta, valonilmaisimesta ja vastaanotin-elektronikasta sekä navigointi- ja inertiajärjestelmistä. Viimeisenä mainittua ei tarvita maasta tehtävissä mittauksissa, koska sen käyttötarkoitus on lentokoneen tai muun ilma-aluksen lentodynamiikan määrittäminen (kallistuminen, kääntyminen ja nyökkäminen).

Tässä työssä käytetty laitteisto (kuva 10) on Geodeettisen laitoksen (FGI engl. Finnish Geospatial Research Institute) kehittämä maailman ensimmäinen hyperspektrinen superjatkumolidar. Laitteisto koostuu valonlähteenä toimivasta superjatkumolaserista (LEUKOS-SM), kollimaattorista (Thorlabs, CFC-5-A), säteenjakajasta, valon keräysoptiikasta, spektrografista (Specim, ImSpector V10), 16-osaisesta vyöryvalodiodimoduulista (First Sensor), jossa on sisäänrakennettu transimpedanssivahvistin (Analog Devices, AD8015), ja 12-bittisestä analogisesta-digitaaliseen konvertterista (SP Devices, ADQ412).(Hakala, Suomalainen et al. 2012)



Kuva 10. Hyperspektrilidarin optinen moduuli. Kollimaattori (A), säteen jakaja (B), parabolinen peili (C), 2D-skanneri (D) ja sensori (E) (Nevalainen, Hakala et al. 2014).

Laitteistossa valonlähteenä on LEUKOS-SM superjatkumolaser, joka tuottaa 1 ns pulsseja, 5 kHz toistotaajuudella, aallonpituuksilla 420-2100 nm. Laserin keskimääräinen ulostuloteho on 40mW. (Nevalainen, Hakala et al. 2014) Laitteisto käyttää mittauksissa kahdeksaa diskreettiä aallonpituuskanavaa, jonka perusteella sitä voisi pitää multispektrisenä. Sen mittauseriaate on kuitenkin hyperspektrinen, koska superjatkumolaserin spektri ja spektrografin ulostulo ovat molemmat jatkuvia. (Hakala, Suomalainen et al. 2012)

Superjatkumolla tarkoitetaan epälineaaristen vuorovaikutusten avulla luotua erittäin laajaspektristä valonkaistaa. Superjatkumo luodaan pumpaamalla valoa laserilla epälineaariin valokuituun, jossa spektrikaistan laajeneminen tapahtuu. Superjatkumon syntymisen taustalla oleva fysiikka vaihtelee riippuen väliaineen kromaattisesta dispersiosta, kuidun pituudesta, pulssin kestosta, pulssin voimakkuudesta ja pumpun aallonpituudesta. Kuitukaapelista poistuva valo sisältää laserille tyypillisiä ominaisuuksia, kuten suuren kirkkouden, tiukasti rajatun säteen ja korkean paikkakoherenssin. Superjatkumovalonlähteen avulla on mahdollista tuottaa valoa, jonka spektri ulottuu ultraviolettialueelta aina lyhytaaltoiselle infrapuna-alueelle asti. Spektroskopian näkökulmasta laajaspektrinen ja hyvin kirkas valonlähde luo monia mahdollisuuksia eri sähkömagneettisten ilmiöiden tutkimiseen. Superjatkumon syntymisestä kiinnostuneen lukijan kannattaa tutustua Dudleyn ja Gentyn et al. 2006 kirjoittamaan artikkeliin. (Dudley, Genty et al. 2006)

### 3 Kiviaineksen laatuun vaikuttavat tekijät

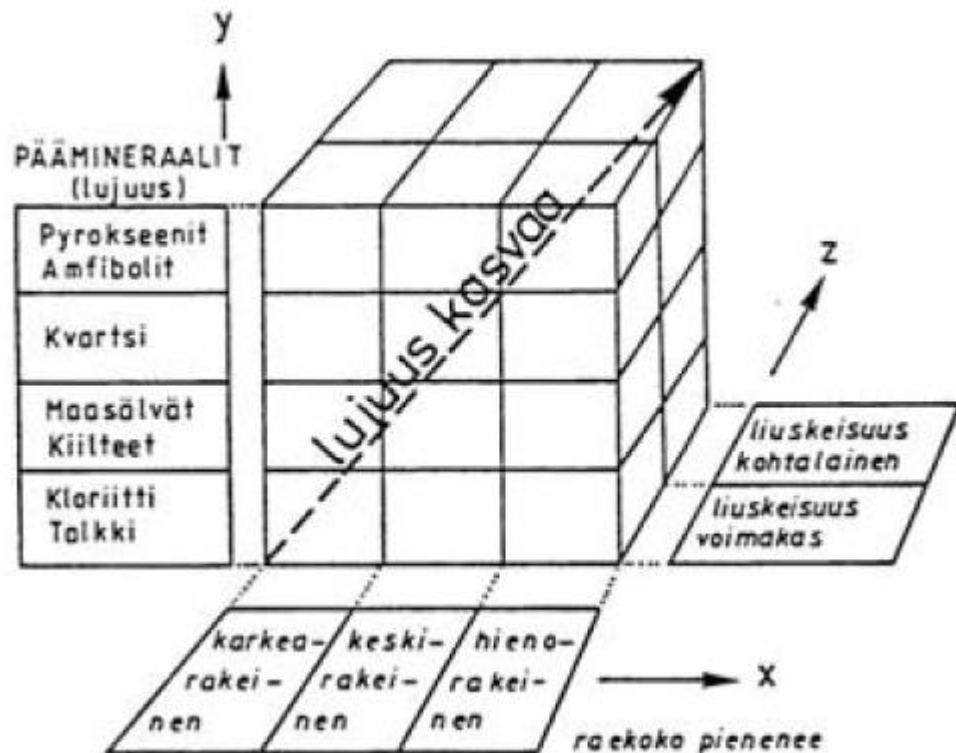
Suomen kallioperä on osa prekambrista Pohjois- ja Itä-Euroopan peruskallioaluetta. Se on muodostunut noin 3000–1400 miljoonaa vuotta sitten. Hallitsevina kivilajeina ovat graniittiset kivilajit, joita arvellaan olevan noin 53 % kallioperästä. Toinen yleinen kivilaji on erilaiset migmatiitit, joita on noin 22 % kallioperästä. Loput koostuvat liuskeista, kvartsiiteista, kalkkikivistä ja emäksisistä magmakivistä. Suomen kallioperä voidaan karkeasti jakaa kahteen vyöhykkeeseen Raahe-Laatokka linjaa pitkin. Linjan pohjoispuolella kallioperä on arkeista yli 2500 miljoonaa vuotta vanhaa, kun taas linjan eteläpuolella kallioperä on varhaisproterotsooista noin 1930–1800 miljoonaa vuotta vanhaa. (Lehtinen, Nurmi et al. 1998) Etelä-Suomen kallioperälle on ominaista, että rapakivigraniitteja ja Satakunnan Jotunisia hiekkakiviä ja diabaaseja lukuun ottamatta, suurin osa siitä on metamorfoitunutta (Eklund 2015).

Kiven syntytytapa vaikuttaa olennaisesti sen lujuusominaisuuksiin. Sedimenttikivet ovat yleensä hauraita ja kestävät huonosti kulutusta. Metamorfiset kivet ovat usein suuntautuneita ja sisältävät paljon pehmeitä kiilteitä. Syväkivet ovat karkearakaisia ja tiiviitä. Vulkaniitit ovat taas hienorakeisia ja huokoisia.

Kiviaineksen laatu koostuu fysikaalisesta lujuudesta ja kemiallisesta kestävydestä. Yleisesti voidaan kuitenkin todeta, että jos kiviaines on lujaa ja kulutuksen kestävä, täyttää se luultavasti suurimman osan vaadituista ehdoista. Lujuutta tarkasteltaessa tulee huomioida sekä kyky kestää naarmutusta (kovuus) että lohkaisua (sitkeys) vastaan. Mineraalien kovutta mitataan yleensä Mohsin asteikolla pehmeimmästä kovimpaan 1-10. Mohsin asteikko on suhteellinen ja lähinnä suuntaa antava. Pehmeät mineraalit ovat Mohsin asteikolla noin 1-3 kun taas kovat mineraalit noin 5-6. Kivet luokitellaan koviksi jos niiden kvartsipitoisuus on yli 40 prosenttia. Pehmeissä kivissä talkin, kiilteen tai kloriitin määrä on yli 40 prosenttia. Pienet määrät pehmeitä mineraaleja vaikuttavat yleensä positiivisesti kiven lujuusominaisuuksiin tehden niistä sitkeitä. Kovat mineraalit kestävät hyvin naarmutusta kun taas sitkeät kestävät hyvin lohkeamista vastaa. Kivet luokitellaan sitkeiksi jos niiden sisältämän amfiboliitin tai pyrokseenin määrä on yli 25 prosenttia. (Nurmikolu 2004)

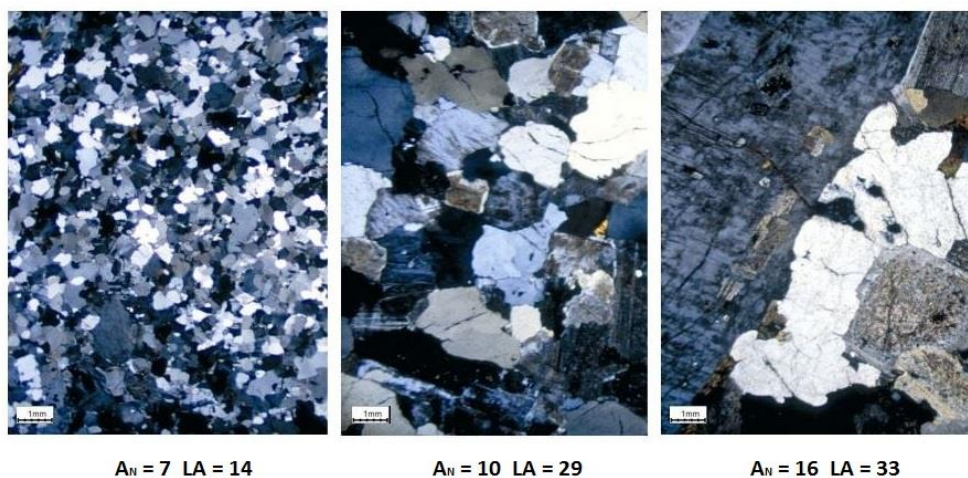
Kiviaineksen lujuuteen vaikuttaa myös kutous, eli se miten mineraalirakeet ovat liittyneet toisiinsa. Kutous määrittelee sen, kuinka hyvin kiviaines lohkeaa rakeiden rajapintoja ja mineraalin sisäisiä lohkosuuntia pitkin. Sälöiset ja epäsäännölliset rakeet muodostavat lujemman kutouksen kuin sileät ja pyöreät rakeet. Myös raekoolla on vaikutusta, koska hienorakeisessa kivilajissa mineraalien välillä on enemmän koheesiovoimia. Kiven liuskeisuudella on myös selkeä vaikutus sen lujuuteen, koska liuskeinen kivi lohkeaa yleensä etevästi liuskeisuuden suuntaan. Anisotrooppisen kiven lujuusparametreissa saattaa olla erittäin suuria eroja heikoimman ja lujimman suunnan välillä. Kiven lujuusparametrit voivat myös muuttua rapautumisen johdosta ja saattavat vaihdella suuresti alueellisestikin. (Nurmikolu 2004) Kiviaineksen lujuutta voidaan arvioida (Kauranne, Gardemeister et al. 1972) mukaisen lujuuskuution avulla (kuva 11). Se ottaa huomioon päämineraalien lujuuden, raekoon ja liuskeisuuden. Se ei kuitenkaan ota huomioon rapautumista.





Kuva 11. Lujuuskuutio (Kauranne, Gardemeister et al. 1972).

Kiviaineksen iskun- ja kulutuksenkestävyys määritellään standardoiduilla menetelmillä. Yleisesti käytössä olevia menetelmiä ovat esimerkiksi Los Angeles -testi (SFS-EN 1097-2, 2010) ja pohjoismainen kuulamyllyarvo (SFS-EN 1097-9 2014). Los Angeles -testi mittaa kiviaineksen iskunkestävyyttä ja sen antamat tulokset kertovat kuinka suuri osa (%) näytteestä läpäisee ennalta määrätyn seulan. Pohjoismainen kuulamyllyarvo mittaa nastarengaskulutuksenkestävyyttä ja sen tulokset ovat kiven kulumisesta johtuvia prosentuaalisia painohävikkejä. Molemmista testeistä suuremmat luvut indikoivat heikommasta kestävyydestä. Kuvassa 12 on nähtävissä raekoon vaikutukset pohjoismaiseen kuulamyllyarvoon ja Los Angeles – lukuun. Sen perusteella on havaittavissa, että raekoolla on suuri merkitys kiviaineksen lujuuteen ja kulutuksenkestävyyteen.



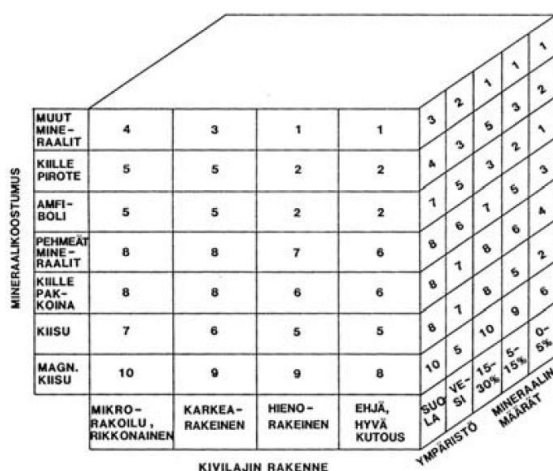
Kuva 12. Raekoon vaikutus pohjoismaiseen kuulamyllyarvoon ja Los Angeles – lukuun (Arnbom 2015).

Kiviaineksen lujuusominaisuuksiin ja kulutuksenkestävyyteen vaikuttavista tekijöistä aina-kin raekoko, mineraalikoostumus ja mineraalien suhteelliset osuudet vaikuttavat heijastus-spektriin (Cloutis 1996, Guha, Rao et al. 2012, Hunt 1977). Tämän perusteella on oletetta-vaa, että kiviainesten lujuusparametrit on johdettavissa niiden spektrisen sormenjäljen pe-rusteella.

Kiviaineksen rapautumista voi tapahtua sekä fysikaalisesti että kemiallisesti. Fysikaalinen rapautuminen johtuu lämpötilan vaihteluista ja kosteudesta. Suomessa yleisin fysikaalista rapautumista aiheuttava tekijä on jäätymis-sulamistoiminta, jonka on havaittu aiheuttavan huomattavaa lujuusominaisuuksien heikkenemistä. Syynä tähän on kiviaineksen laajentunut huokostilavuus ja ominaispinta-ala, joiden johdosta gravitaatiohuokosiin virtaavan veden määrä kasvaa ja altistaa pakkasrapautumiselle. (Nurmikolu 2004)

Kemiallinen rapautuminen johtuu ilmakehän, kiven ja veden välisistä kemiallisista reakti-oista, joissa kiven mineraalikoostumus muuttuu. Rapautumisnopeuteen vaikuttaa olennai-sesti liuottamistoimintaan saatavilla olevan veden määrä, rakotiheys, huokoisuus ja vir-tausolosuhteet. Yleisiä kemiallisen rapautumisen prosesseja ovat liukeneminen, hapettumi-nen, pelkistyminen, karbonaatio, hydraatio, hydrolyysi, fiksaatio ja kelaatio. Kemialliseen rapautumiseen vaikuttavat ympäristön pH, hapetus – pelkistys -olosuhteet, vaikuttava pinta-ala ja mineralogia. Rapautumisnopeus on kalliokiviaineksestä tehdyissä murskeissa huomattavasti avoimia kalliopaljastumia suurempi, suuremmasta ominaispinta-alasta johtuen. (Nur-mikolu 2004)

Kivimineraalien kemiallisen rapautumisalttiuden ja niiden kiteytymisjärjestyksen välillä on havaittu olevan korrelaatiota siten, että varhaisimmassa vaiheessa magmasta kiteytyvät mi-neraalit rapautuvat herkemmin kuin myöhemmässä vaiheessa ja matalammassa lämpöti-lassa kiteytyvät mineraalit. Mineraalien rapautumispotentiaali-indeksit (WPI, engl. weathe-rin potential index) seuraavatkin hyvin Bowenin reaktiosarjaa, joka kuvaa mineraalien ki-teytymisjärjestystä. Päällystekivien rapautumisherkkyyttä voidaan arvioida (Nieminen, Jä-niskangas 1990) kehittämän rapautumisherkkyysskuution avulla (kuva 13), jossa huomioi-daan mineraalikoostumus, kivilajin rakenne, ympäristö ja mineraalien määrät.



Kuva 13. Päällystekivien rapautumisherkkyy (Nieminen & Jäniskangas).

Koska rapautumisherkkyyteen vaikuttavista tekijöistä ainakin mineraalikoostumus ja mineraalien määrät sekä raekoko ovat heijastusspektriin vaikuttavia ominaisuuksia (Cloutis 1996, Guha, Rao et al. 2012, Hunt 1977), voidaan rapautumisherkkyyden arvioimista hyperspektRIDATAN avulla pitää mahdollisena.



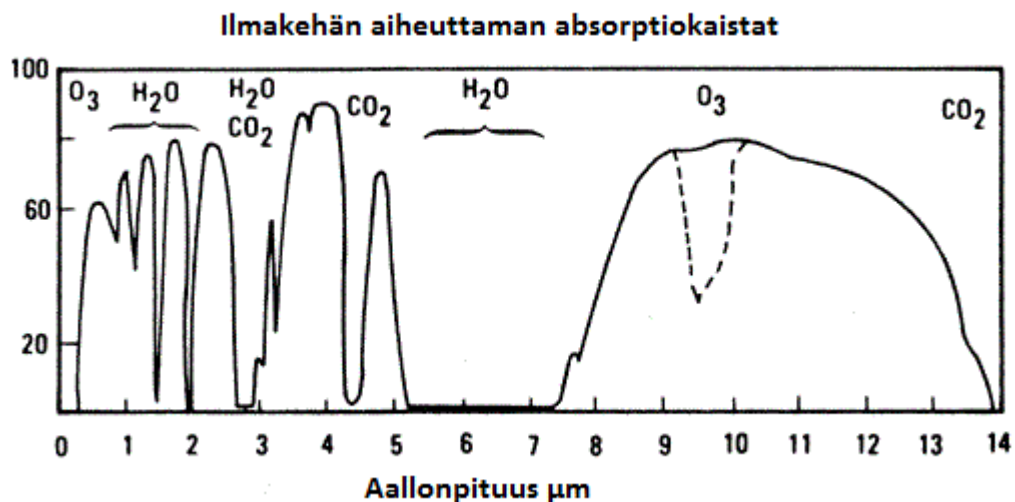
## 4 Hyperspektridatan käsittely

### 4.1 Datan korjaukset

Kentällä kuvattuun hyperspektriseen dataan joudutaan yleensä tekemään korjauksia, jotka johtuvat ilmakehän vaikutuksesta, mittauskulmasta, sääolosuhteista tai mittausetäisyydestä. Mitattu säteilyn intensiteetti ei vastaa kohteen heijastuskykyä eli reflektanssia, vaan se on osittain vallitsevien olosuhteiden muokkaama. Jotta kohteen identifioinnin kannalta olennainen reflektanssitieto voidaan saada, tulee mittausdataan tehdä korjauksia, jotka kompensoivat vallitsevien olosuhteiden vaikutuksia. Aktiivisen hyperspektrikuvantamisen etuna on, että mittauskulman muutosten tai sääolosuhteiden aiheuttama vaikutus jää passiivista hyperspektrikuvantamista pienemmäksi. Koska valonlähde- ja vastaanotinoptiikka sijaitsevat samassa pisteessä, ei mittauskulman vaikutus mitatun spektrin intensiteettiin ole yhtä merkittävä kuin passiivisessa kuvantamisessa. (Nischán, Joseph et al. 2003) Lyhyillä etäisyyksillä tehdyissä kenttämittauksissa ilmakehän vaikutuksia yhden kuvan alueella voidaan pitää vakioina.

#### 4.1.1 Ilmakehäkorjaus

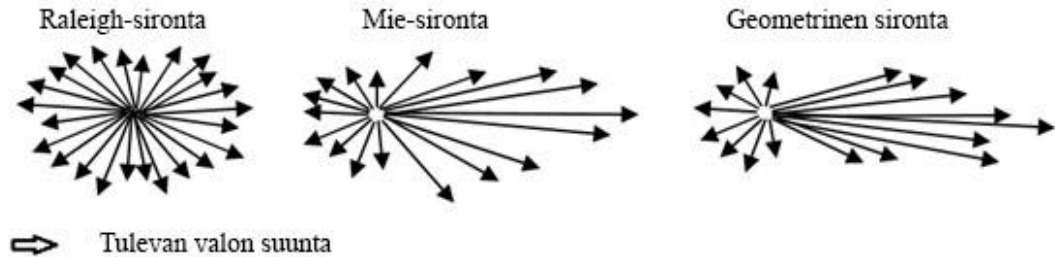
Ilmakehäkorjauksen tarkoituksena on poistaa ilmakehän kaasujen ja hiukkasten aiheuttamat vaikutukset mittausdatasta. Ilmakehän kaasut ja hiukkaset aiheuttavat muutoksia sen läpi kulkevan säteilyn taajuuteen, suuntaan ja intensiteettiin. Merkittävimmät tekijät ilmakehän aiheuttamassa absorptiossa ovat vesihöyry, hiilidioksidi, otsoni, typen oksidit, häikä, metaani ja happi. (Gao, Montes et al. 2009) Edellä mainituista tekijöistä merkittävin on yleensä vesihöyry, joka aiheuttaa syvät absorptiokaistat 1400 nm ja 1900 nm aallonpituuksilla (kuva 14). Absorptiokaistojen syvyys on myös suhteessa mittausetäisyyteen, joten mittausetäisyyden kasvaessa absorptiokaistat syventyvät.



Kuva 14. Ilmakehän aiheuttamat absorptiokaistat (Sabins Jr 1978).

Ilmakehässä olevat aerosolit ja hiukkaset aiheuttavat myös sirontaa, jolloin fotonit eivät absorboitu vaan muuttavat suuntaansa. Sironta voi tapahtua joko Rayleigh-, Mie- tai geometrisena

sirontana (kuva 15). Rayleigh-sironta tapahtuu partikkeleista jotka ovat huomattavasti säteilyn aallonpituutta pienempiä, Mie-sironta tapahtuu partikkeleista jotka ovat suurin piirtein yhtä suuria kuin läpikulkevan säteilyn aallonpituus ja geometrinen sironta (valikoimaton sironta) tapahtuu selkeästi säteilyn aallonpituutta suuremmista partikkeleista. Sironta voi tapahtua myös epäelastisena sirontana, kuten Raman-sirontana, mutta se huomattavasti epätodennäköisempää ja näin ollen intensiteetiltään erittäin alhaista. (Kumar 2005)



Kuva 15. Eri sirontatyypit (Nave 2016).

Ilmakehän aiheuttamat muutokset heijastusspektrissä voidaan kompensoida joko mallintamalla tai mittaamalla vallitsevat olosuhteet, tai käyttämällä kalibroitipaneelleja, joiden heijastavuusominaisuudet tunnetaan. Tässä työssä tehdyissä mittauksissa käytettiin Spectralon-paneelia (99 %) intensiteetin muuttamiseksi reflektanssiksi.

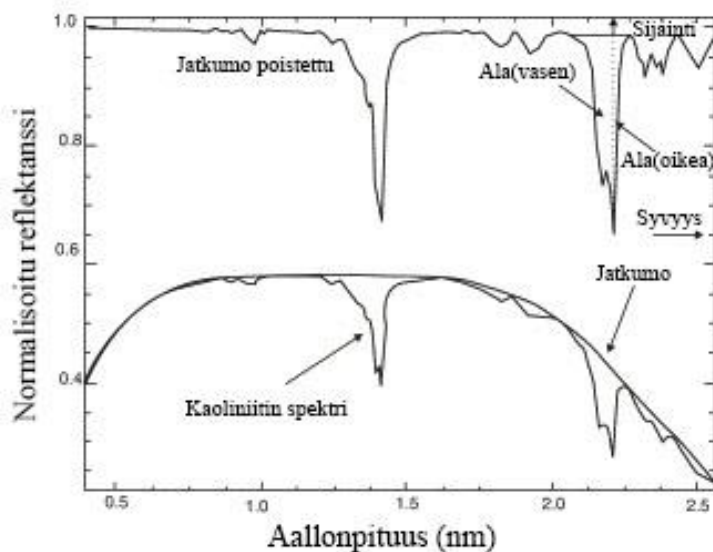
#### 4.1.2 Jatkumon poistaminen

Jatkumo (engl. continuum) on spektrin suurimpien heijastavuuspisteiden kautta kulkeva kuvitteellinen käyrä (kuva 16). Koska eri laitteistoilla ja eri ajankohtina tuotetut spektrit eroavat toisistaan, eivät ne yleensä ole suoraan vertailukelpoisia keskenään. Spektritiedon kannalta olennaisimmat absorptiokaistat jäävät jatkumon poistamisessa ennalleen, mutta eri mittaustekniikoiden aiheuttamat taustaheijastuvuuden intensiteettierot saadaan poistettua. Poistamalla jatkumo spektristä saadaan spektri normalisoitua, jolloin siihen voidaan käyttää Spectral Feature Fittingin (SFF) kaltaisia absorptiopiirteisiin perustuvia menetelmiä.

Normalisoidussa spektrissä jatkumon kanssa yhdenmukaiset pisteet saavat arvon 1 ja muut pisteet arvon  $< 1$ . Jatkumo poistetaan jakamalla alkuperäisen spektripisteen arvo jatkumon arvolla:

$$S_{\alpha} = S/C, \quad (10)$$

missä  $S_{\alpha}$  on jatkumopoistettuspektri,  $S$  on alkuperäinen spektri ja  $C$  on jatkumo. (Clark, Swayze et al. 2003)



Kuva 16. Jatkumon poistaminen spektristä (van der Meer 2004).

Jatkumon poistoa ei käytetty tässä työssä tehdyissä analyyseissa, koska kaikki mittaukset tehtiin samalla laitteistolla ja samoissa olosuhteissa.

## 4.2 Piirteiden vähentäminen ja muuntaminen

Piirteiden valinnan tarkoituksena on luokittelun kannalta tärkeimpien piirteiden valinta niin, että mahdollisimman suuri osa luokittelun kannalta olennaista informaatiota saadaan säilytettyä. (Theodoridis, Koutroumbas 2009) Hyperspektrikuvien sisältämä suuri informaation määrä tarjoaa mahdollisuuksia useiden eri luokkien tunnistamiseksi, mutta se sisältää myös ongelmia. Suuri datan määrä vaatii paljon laskentatehoa ja suuri dimensioiden määrä vaatii suuret harjoitussarjat tuottaakseen tarkkaa luokittelua (Serpico, D'Inca et al. 2003). Jos datan dimensioiden kasvaessa harjoitussarjan koko ei kasvaa samaa vauhtia, laskee luokittelun tarkkuus Hughesin efektin johdosta (Hughes 1968). Koska hyperspektrikuvien sisältämä data on erittäin kapeilta ja lähekkäisiltä aallonpituusalueilta hankittua, korreloivat vierekkäiset kanavat yleensä hyvin vahvasti keskenään.

Alkuperäinen piirrejoukko  $D$  vastaa kuvattujen kanavien määrää. Piirteiden valinnalla tai poistamisella voidaan luoda uusi piirreavaruus ( $d < D$ ), joka sisältää suurimman osa alkupe-  
räisestä informaatiosta. Piirteiden vähentämisen myötä tarve erittäin suurille harjoitussarjoille vähenee ja laskentatehoa ei tarvita yhtä paljon.

Seuraavaksi tässä työssä esitellään yleisesti käytössä olevia piirteiden valintamenetelmiä. Koska tämän työn yhteydessä tehdyissä mittauksissa käytetty laitteisto pystyi tuottamaan ainoastaan kahdeksan spektrikanavaa, ei piirteiden vähentämisellä ollut tarvetta. Sen sijaan pääkomponenttimuunnosta käytettiin muuttujien vähentämisen vaikutusten tarkastelussa.

### 4.2.1 Sequential Forward Selection

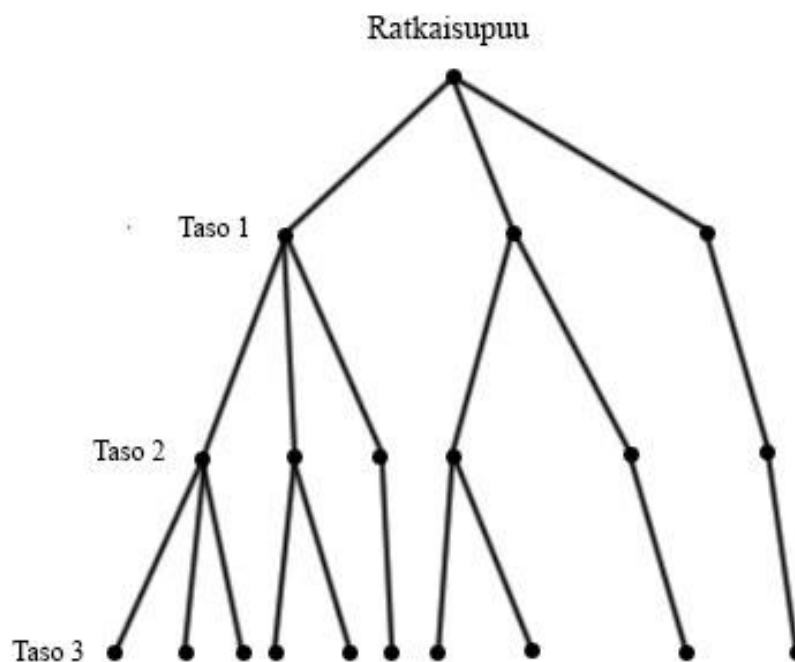
Sequential Forward Selection (SFS) on piirteiden valintamenetelmä, jossa aloitetaan tyhjästä piirrejoukosta johon lisätään aina yksi, suurimman kriteerifunktion arvon antava piirre. Piirteiden lisäämistä jatketaan, kunnes uusien piirteiden lisääminen ei enää paranna piirrejoukon kuvaamista. SFS:n heikkoutena on, että kertaalleen lisättyä piirrettä ei enää pysty poistamaan, vaikka se osoittautuisi tarpeettomaksi myöhemmin lisättyjen piirteiden jälkeen. (Serpico, Bruzzone 2001)

### 4.2.2 Sequential Backward Selection

Sequential Backward Selection (SBS) on SFS:n kaltainen piirteiden valintamenetelmä, jossa aloitetaan täydestä piirrejoukosta ja poistetaan yksi piirre kerrallaan, kunnes haluttu piirrejoukko on saavutettu. SBS:ssä poistetaan aina se piirre, jonka poistamisen jälkeen kriteerifunktion arvo on mahdollisimman suuri. SBS on SFS:ää raskaampi algoritmi, koska ensimmäisiä piirteitä poistettaessa kriteerifunktiot joudutaan laskemaan hyvin moniulotteisessa piirreavaruudessa ( $d = D-1$ ). SFS:n tapaan, SBS ei kykene jälkeinpäin tekemään muutoksia valittuun piirrejoukkoon, eli jo poistettua piirrettä ei voida enää lisätä piirrejoukkoon. (Serpico, Bruzzone 2001)

### 4.2.3 Branch and bound

Branch and Bound (BB) on (Narendra, Fukunaga 1977) kehittämä piirteiden valintamenetelmä, jossa käydään läpi kaikki piirteiden alijoukot. Todellisuudessa algoritmi ei suorita täydellistä tarkastelua, vaan alitettuaan kriteerifunktion kynnysarvon se palaa takaisin lähimpää solmukohtaan, jota ei ole vielä tutkittu. Algoritmi käy piirteitä läpi ratkaisupuu-mallina, jossa tarkastelu aloitetaan juuresta, jossa yhtään piirrettä ei ole vielä poistettu. Piirteitä poistetaan jokaisessa solmukohdassa, kunnes halutun kokoinen osajoukko on saavutettu. Kriteerifunktion arvo tälle osajoukolle asetetaan kynnysarvoksi ja palataan takaisin lähimpään haarautumiskohtaan ja tutkitaan tämän osajoukon arvoa. Jos kynnysarvo alittuu, palaa funktio takaisin edelliseen haarautumiskohtaan. Jos taas kynnysarvo ylittyy, asetetaan tämän haaran antama arvo uudeksi kynnysarvoksi. Tätä jatketaan, kunnes koko puu on käyty läpi. (Heikkilä, Törmä 2003) Branch and Bound menetelmää havainnollistetaan kuvassa 17.



Kuva 17. Branch and Bound –ratkaisupuu-malli.

#### 4.2.4 Plus l - take away r

Plus l – take away r -algoritmi on SFS ja SBS – algoritmien yhdistelmä, joka ensin valitsee l piirrettä SFS:n avulla ja sen jälkeen poistaa niistä r piirrettä SBS:n avulla. Arvot l ja r on käyttäjän määriteltävä itse ennen algoritmin ajamista. Menetelmän heikkoutena on SFS ja SBS -algoritmeihin liittyvä piirteiden keskinäisestä riippuvuudesta piittaamattomuus. Tämä ongelma voidaan poistaa käyttämällä yleistettyä plus l – take away r -algoritmia. (Heikkilä, Törmä 2003, Serpico, Bruzzone 2001) Toinen plus l – take away r -algoritmin ongelma on, että parametrien l ja r valintaan ei ole olemassa mitään teoreettista kriteeriä. (Serpico, Bruzzone 2001)

#### 4.2.5 Pääkomponenttimuunnos

Pääkomponenttimuunnos, josta käytetään myös nimityksiä Hotelling-muunnos ja Karhunen-Löve -muunnos on yleisesti käytetty tilastollinen menetelmä, jonka avulla alkuperäiset muuttujat esitetään uusina korreloimattomina muuttujina. Pääkomponenttimuunnos ei varsinaisesti ole piirteiden valintamenetelmä, koska se ei poista osaa alkuperäisistä piirteistä, vaan muuntaa ne suurimman varianssin suuntaisiksi uusiksi piirteiksi. Saatuja uusia muuttujia kutsutaan pääkomponenteiksi ja ne ovat alkuperäisten muuttujien lineaarikombinaatioita. Pääkomponenttianalyysillä pyritään yleensä vähentämään datan dimensioita, selittämään mittaustulosten taustalla piileviä latenttimuuttujia, paljastamaan ryhmittymiä ja tunnistamaan datan ulkopuolelle jääviä näytteitä. Pääkomponentteja voi olla korkeintaan yhtä paljon kuin alkuperäisiä muuttujia. (Jolliffe 2002)

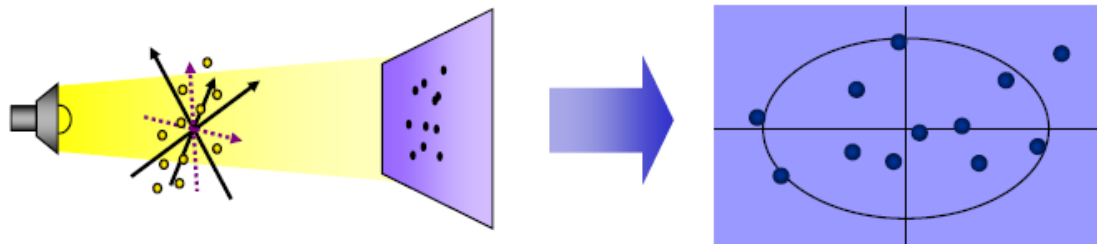
Ennen pääkomponenttimuunnoksen tekemistä data yleensä joko nollakeskiarvoistetaan tai autoskaalataan. Nollakeskiarvoistamisessa datasta vähennetään sen keskiarvo. Tämä tehdään yleensä jokaiselle sarakkeelle (aallonpituudelle) erikseen. Nollakeskiarvoistamisen voi ajatella datapisteiden keskiarvon siirtämisenä origoon. Autoskaalauksessa datasta vähennetään sen keskiarvo ja jaetaan keskihajonnalla. Näin ollen sarakkeiden varianssiksi tulee 1. (Jolliffe 2002)

Pääkomponenttimuunnoksessa ensimmäinen pääkomponentti valitaan niin, että havaintopisteiden välinen varianssi on mahdollisimman suuri. Se siis esittää maksimi varianssin suunnan datassa. Varianssin maksimi löytyy datan kovarianssimatriisin suurinta ominaisarvoa vastaavasta ominaisvektorista. Pääkomponentit lasketaan niitä vastaavan ominaisvektorin ja muuttujien pistetulona:

$$PC_i = \boldsymbol{\beta} \cdot \mathbf{x} = \beta_i * x_i + \beta_j * x_j + \dots + \beta_n * x_n, \quad (11)$$

missä  $\boldsymbol{\beta}$  on ominaisvektorin alkio ja  $\mathbf{x}$  on muuttujavektori.

Ensimmäinen pääkomponentti kulkee datan keskiarvon läpi. Toinen pääkomponentti valitaan niin, että se on kohtisuorassa ensimmäistä pääkomponenttia vastaan ja se selittää suurimman osan jäännöksen varianssista ja sen suunnan. Myös toinen pääkomponentti kulkee datan keskiarvopisteen läpi. Ensimmäiset kaksi pääkomponenttia virittävät tason, johon jokainen n-ulotteinen datapiste voidaan projisoida (kuva 18). Pääkomponentit ovat ortogonaalisia ja näin ollen keskenään korreloimattomia. (Jolliffe 2002)



Kuva 18. Pääkomponenttien projektio tasolle (Wiklund 2008)

Pääkomponenttianalyysin tuloksia voidaan tulkita pistemäärien (scores) ja latausten (loadings) avulla. Pistemäärä kuvailee havaintoja ja mahdollistaa niiden tarkastelun projektiotasolla. Lataukset kuvailevat muuttujia ja selittävät havaintojen sijoittumisen pistemääräkuvaajassa (score plot). Pistemäärät (havainnot) saattavat muodostaa ryhmittymiä, joiden sisällä olevat havainnot sisältävät yhteneväisiä ominaisuuksia. Tarkastelemalla latauksia voidaan löytää muuttujat, jotka aiheuttavat edellä mainitun ryhmittymisen. (Jolliffe 2002)

Hyperspektrikuvat saattavat sisältää satoja kanavia, jotka osittain korreloivat vahvasti keskenään. Edellä mainitun kaltaisessa tilanteessa dimensioiden vähentäminen pääkomponenttimuunnoksen avulla saattaa nopeuttaa laskenta-aikaa merkittävästi ja mahdollistaa havaintojen visuaalisen tulkinnan.

Tässä työssä tehtyjen mittausten analysoinnissa pääkomponenttimuunnosta käytettiin dimensioiden vähentämisen vaikutusten tarkastelussa.

### 4.3 Luokittelualgoritmit

Hyperspektridatan käytössä merkittävä tekijä on käyttötarkoitukseen sopivan luokittelualgoritmin valinta. Luokittelualgoritmin tarkoituksena on verrata mitattua dataa spektrikirjastosta tai itse kuvasta irrotettuun referenssidataan. Luokittelualgoritmit etsivät yhtäläisyyksiä referenssispektrin ja mitatun spektrin välillä. Oikean luokittelualgoritmin valinta perustuu usein eri menetelmien testaamiseen tai kokemusperäisen tiedon käyttämiseen. Luokittelualgoritmin valinnalla voidaan vaikuttaa myös esimerkiksi valaistusolosuhteiden, mittaustäi-syyden tai mittauskulman vaikutuksiin luokittelussa. Seuraavaksi esitellään yleisesti käytettyjä luokittelualgoritmeja, joista osaa on käytetty tässä työssä tehtyjä mittauksia analysoidessa.

#### 4.3.1 Spectral Unmixing

Hyperspektrikuvantamisessa kohdataan usein tilanteita, joissa yksittäinen pikseli ei edusta ainoastaan yhtä luokkaa (esim. graniittia tai kalimaasälpää), vaan se koostuu useista päätejäsenistä. Jos yhden pikselin alueella on jokin selkeästi hallitseva päätejäsen, tehdään luokittelu yleensä sen mukaisesti. Jos hallitsevaa päätejäsentä ei löydy, voidaan kohteen luokitteluun käyttää Spectral Unmixingia. Sekoittuneen pikselin spektri voidaan kuvata usean päätejäsenen lineaarikombinaationa. Referenssispektrejä hyväksi käyttäen voidaan sekoittuneiden päätejäsenten pitoisuudet kuvatussa spektrissä selvittää. Lineaarinen Spectral Unmixing – malli (Tai LMM eng. Linear Mixing Model) on muotoa:

$$\mathbf{y} = \sum_{r=1}^R \mathbf{m}_r \alpha_r + \mathbf{n}, \quad (12)$$

missä  $\mathbf{y} = [y_1, \dots, y_L]^T$  on sekoittuneen pikselin spektri, jossa  $\mathbf{m}_r$  r:nnen päätejäsenen spektri ja  $\alpha_r$  on sen suhteellinen osuus sekoituksessa ja  $\mathbf{n}$  on kohina. (Dobigeon, Tournier et al. 2008)

Kappaleesta 2.1.2 löytyy esimerkki spektrikirjastosta haettujen referenssispektrien avulla rakennetusta lineaarisesta kompositiospektristä.

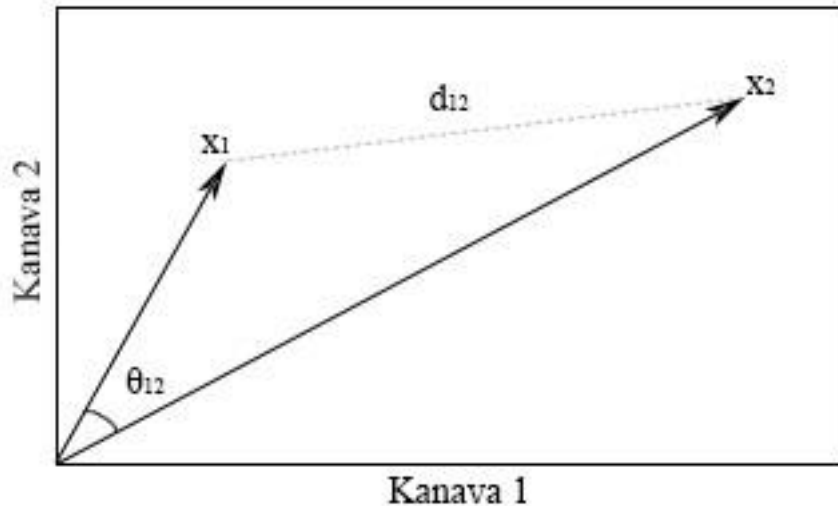
#### 4.3.2 Spectral Feature Fitting

Spectral feature fittingissä (SFF) referenssispektristä ja mitatusta spektristä poistetaan jatkumo (kappale 4.1.2) ja tarkastelun kohteena ovat jäljelle jäävät absorptiopiirteet. Referenssispektrin ja mitatun spektrin yhtäläisyyttä verrataan pienimmän neliösumman menetelmällä. (Clark, Gallagher et al. 1990) SFF on erityisen hyvä silloin kun referenssispektri ja mitattu spektri ovat erilaisilla mittausten menetelmillä hankittuja. Jatkumon poistamisella eliminoidaan kokonaisintensiteetin eroavaisuudet, jolloin eri mittausten menetelmillä hankitut spektrit ovat paremmin vertailtavissa.

Spectral Feature Fittingiä ei käytetty tässä työssä tehdyissä analyyseissa, koska kaikki mitaukset tehtiin samalla laitteistolla ja samoissa olosuhteissa.

### 4.3.3 Spektrikulman mukainen luokittelu

Spektrikulman mukaisessa luokittelussa käytetään yleensä Spectral Angle Mapper-algoritmitmia, joka perustuu mitattujen ja tunnettujen pikselivektoreiden välisen kulman vertailuun. Kuten kappaleessa 2.1.3 todettiin, hyperspektrikuvan pikseleitä voidaan käsitellä vektoreina, joilla on suunta ja pituus. Pikselivektoreita voidaan siis verrata joko referenssikirjastoista tai itse kuvasta erotettuihin referenssivektoreihin. Kuvasta 19 nähdään spektrikulman ( $\theta_{12}$ ) ja kappaleessa 4.3.5 esiteltävän etäisyyden ( $d_{12}$ ) laskentaperiaate.



Kuva 19. Spektrikulman ( $\theta_{12}$ ) ja etäisyyden ( $d_{12}$ ) laskentaperiaate.

Spectral Angle Mapper (SAM) ei ota kantaa pikseleiden väliseen intensiteettieroon (vektoreiden pituuseroon), vaan ainoastaan niiden väliseen kulmaan. Vektoreiden välinen kulma kuvaa spektristen ominaisuuksien vastaavuutta referenssidatan ja mitatun datan välillä. Vektoreiden välisen kulman funktio voidaan johtaa pistetulon määritelmästä:

$$\mathbf{t} \cdot \mathbf{r} = \|\mathbf{t}\| \|\mathbf{r}\| \cos \theta = \mathbf{t}^T \mathbf{r}, \quad (13)$$

missä  $\mathbf{t}$  on havaintovektori ja  $\mathbf{r}$  on referenssivektori. Vektoreiden väliseksi kulmaksi ( $\theta$ ) saadaan:

$$\theta = \cos^{-1} \left( \frac{\mathbf{t} \cdot \mathbf{r}}{\|\mathbf{t}\| \|\mathbf{r}\|} \right), \quad (14)$$

joka voidaan myös ilmaista summafunktioiden avulla:

$$\theta = \cos^{-1} \left( \frac{\sum_{i=1}^{nb} t_i r_i}{\sqrt{\sum_{i=1}^{nb} t_i^2} \sqrt{\sum_{i=1}^{nb} r_i^2}} \right), \quad (15)$$

missä  $nb$  on kanavien lukumäärä,  $t$  on havaittu spektri ja  $r$  on referenssispektri (Yuhás, Goetz et al. 1992). SAM – luokittelun tuloksena syntyy kuva, jossa tummat pikselit ( $\theta$  on pieni) kuvaavat hyvää vastaavuutta referenssispektrin ja mitatun datan välillä. Kuvan visuaalisuuden parantamiseksi värit käännetään usein ympäri, jolloin vaaleat pikselit vastaavat samankaltaisuutta.



Tässä työssä käytetty spektrikulman luokittelu tehtiin SAM – algoritmilla, korvaten havaittu spektri  $t$ , kahden sadan havaitun pisteen keskiarvolla. Referenssispektrinä ( $r$ ) käytettiin vastaavanlaisten keskiarvospektrien keskiarvoja (harjoitussarjaluokkien keskiarvoja).

#### 4.3.4 Korrelaatioluokittelu

Korrelaation mukaisessa luokittelussa käytetään Spectral Correlation Mapper (SCM) nimistä algoritmia, joka on De Carvalho ja Menesesin (2000) esittämä kehittyneempi versio SAM:ista, jossa havaittu- ja referenssispektri standardoidaan vähentämällä jokaisesta spektrivektorista niiden keskiarvo:

$$SCM = \frac{\sum_{i=1}^{nb} (t_i - \bar{t})(r_i - \bar{r})}{\sqrt{\sum_{i=1}^{nb} (t_i - \bar{t})^2} \sqrt{\sum_{i=1}^{nb} (r_i - \bar{r})^2}}, \quad (16)$$

missä  $\bar{t}$  on havaintospektrin kaikkien kanavien keskiarvo ja  $\bar{r}$  on referenssispektrin kaikkien kanavien keskiarvo. SCM eroaa olennaisesti SAM:ista siinä, että se voi saada arvoja väliltä -1..1, kun taas SAM saa arvoja väliltä 0-90° (tai  $\cos \theta = 0..1$ ). SCM siis erottaa myös negatiivisen korrelaation. (De Carvalho, Meneses 2000)

Tässä työssä käytetty korrelaatioluokittelu oli SCM:ia vastaava Pearsonin korrelaatiokerroin, joka laskettiin referenssispektien keskiarvon ja tuntemattomien havaintojen välille (katso kappale 5.4).

#### 4.3.5 Etäisyysluokittelu

Etäisyysluokittelu tehdään käsittelemällä havainto- ja referenssispektrit  $n$ -ulotteisina vektoreina ja laskemalla niiden välinen euklidinen etäisyys ( $D$ ). Euklidisen etäisyyden laskentaa havainnollistaa kuva 4.5.

$$D = \sqrt{(t_1 - r_1)^2 + \dots + (t_n - r_n)^2}, \quad (17)$$

missä  $t$  on havaittu spektri ja  $r$  on referenssispektri. Etäisyysluokittelu voidaan tehdä jonkin raja-arvon perusteella tai luokittelemalla näyte lähimpään mahdolliseen luokkaan. (Campbell, Wynne 2011) Etäisyysluokittelun heikkoutena ovat valaistusolosuhteiden muutokset, jotka vaikuttavat selvästi sen antamiin tuloksiin. Tässä työssä tehdyt etäisyysluokittelut on tehty lähimmän luokan periaatteella.

#### 4.3.6 Ryhmittelyanalyysi

Ryhmittelyanalyysi (klusterianalyysi) on menetelmä, jonka avulla on mahdollista jakaa suuri joukko havaintoja pienempiin kokonaisuuksiin. Ryhmittelyanalyysi sopii havaintojen luokitteluun, eikä se tarvitse välttämättä tietoa luokkien määrästä tai luokkien keskipisteistä. Jos ennakkotietoa edellä mainituista arvoista ei ole, voidaan sopiva luokkien määrä päätellä dendrogrammin avulla tai kokeilemalla. (OriginLab 2015)

Hierarkkisessa ryhmittelyanalyysissä käytetään usein niin sanottua kokoavaa menetelmää, jossa alkuvaiheessa kaikki havainnot ovat erillisissä ryhmissä ja ryhmiä yhdistellään valitun kriteerifunktion mukaisesti, kunnes jäljellä on haluttu määrä ryhmiä. Ryhmittelyanalyysin tuloksiin vaikuttavat etäisyyksien laskemiseen ja ryhmittelemiseen valitut menetelmät. (Everitt, Landau et al. 2011) Tässä työssä tehtyjen analyysien laskentaan käytettiin euklidista etäisyyttä ja Wardin menetelmää. Wardin menetelmän hyvä puoli on, että se pyrkii vähentämään ryhmän sisäisten etäisyyksien neliöiden summaa. Huonona puolena on se, että ohjaamattomana sillä on taipumus tehdä ryhmittymistä hyvin pieniä. (OriginLab 2015)

Koska hierarkkinen ryhmittelyanalyysi on raskas analysointimenetelmä, soveltuu se huonosti isoille aineistoille. Nopeampi menetelmä on käyttää k-keskiarvo -ryhmittelyanalyysiä, joka on laskennallisesti huomattavasti kevyempi. K-keskiarvo analyysi jakaa n havaintoa k:hon luokkaan niiden keskiarvojen etäisyyksien perusteella. K-keskiarvo analyysi luokittelee havainnot ennalta määrättyihin luokkiin, joten se tarvitsee lähtötiedoiksi joko ryhmien keskipisteet tai niiden määrän. Yksi tapa tehdä ryhmittelyanalyysiä on selvittää ryhmien keskipisteet edustavien harjoitussarjojen ja hierarkkisen ryhmittelyanalyysin avulla ja käyttää niitä alkuarvoina k-keskiarvo analyysissä. (OriginLab 2015) Tässä työssä tehdyt analyysit tehtiin edellä mainitulla menetelmällä.

#### 4.3.7 Kanoninen erotteluanalyysi

Erotteluanalyysin tarkoituksena on löytää tunnettuja ryhmiä mahdollisimman hyvin erotteleva sääntö. Kanoninen erotteluanalyysi toimii pääkomponenttianalyysin tapaan, luoden uusia, niin sanottuja kanonisia muuttujia, jotka ovat alkuperäisten muuttujien lineaarikombinaatioita. Toisin kuin pääkomponenttianalyysissä, kanonisessa erotteluanalyysissä uudet muuttujat pyrkivät selittämään ryhmien sisäisen varianssin eikä koko joukon varianssia. Toinen merkittävä ero pääkomponenttianalyysin ja kanonisen erotteluanalyysin välillä on se, että kanonisessa erotteluanalyysissä tarvitaan harjoitussarjoja luokkien määrittämistä varten. Kanonista erotteluanalyysiä ei siis voida täysin automatisoida, mutta hyvien harjoitussarjojen rakentaminen saattaa lisätä erottelutarkkuutta merkittävästi. (Guang, Maclean 2000)

Kanonisen erotteluanalyysin tarkoituksen on löytää alkuperäisten muuttujien ( $x$ ) lineaarikombinaatioita ( $l_1 \dots l_k$ ), jotka erottelevat ennalta määrätty luokat mahdollisimman hyvin. Luokkien välinen erottuvuus on määritelty luokkien ( $y$ ) sisäisten keskiarvojen erotuksen itseisarvona jaettuna niiden keskihajontojen ( $s$ ) summalla. Kahden luokan tapauksessa erottuvuudelle saadaan:

$$Erottuvuus = \frac{|\bar{y}_1 - \bar{y}_2|}{s_{y1} + s_{y2}}, \quad (18)$$

Kaavasta  $x$  voidaan nähdä, että erottuvuus on sitä parempaa, mitä suurempi luokkien keskiarvojen erotus on ja mitä pienempi luokkien keskihajontojen summa on. Kun huomioidaan, että  $y = l'x$ , yleistetyn tapauksen erottuvuuden maksimille saadaan:

$$\frac{l' \left[ \sum_{i=1}^g (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x}) \right] l}{l' \left[ \sum_{i=1}^g \sum_{j=1}^{n_i} (\bar{x}_{ij} - \bar{x})(\bar{x}_{ij} - \bar{x}) \right] l} = \frac{l' M l}{l' N l}, \quad (19)$$

missä  $\bar{x}_i$  on keskiarvo vektori,  $\bar{x}$  on kokonaiskeskiarvo (skalaari),  $g$  on luokkien määrä ja  $n$  on muuttujien määrä. Vektorit  $l$ , jotka maksimoivat erottuvuuden, ovat matriisin  $N^{-1}M$  standardoituja ominaisvektoreita  $l'_i = e_i$ . Kanonisten erottelufunktioiden ( $l'_k x = e_k x$ ) määrä on yleensä yhtä suuri, kuin matriisin  $N^{-1}M$  positiivisten ominaisarvojen määrä. (Guang, Maclean 2000)

Kun luokat ja niiden kanoniset kertoimet on luotu harjoitussarjojen avulla, voidaan tuntemattomia näytteitä verrata malliin ja luokitella lähimpään mahdolliseen luokkaan.

## 5 Näytteet ja menetelmät

### 5.1 Näytteet

Seuraavaksi tässä työssä esitellään mittauksissa käytetyt kiviläytteenä. Kivien mineraalikoostumus on valmistajan ilmoittama ja (SFS-EN 12407 2009) -standardin mukainen. Tarkemmat valokuvat näytteistä löytyvät liitteestä 1. Edellä mainitut kuvat on otettu analyysiin valituista kiviläytteenä. Kiviläytteenä numerointi oli tehty jo ennen tätä tutkimusta, joten näytteenä alkuperäinen numerointi säilytettiin.

#### **Amadeus, Sulkavan monivärinen gneissi**

Amadeus on Etelä-Savosta, Sulkavalta louhittu metamorfinen granaattikordieriittigneissi. Amadeus on suuntautunut, hieno- ja keskirakeinen kivi, jonka päävärit ovat musta ja harmaa. Lisäksi kivessä on havaittavissa punertavia ja vihertäviä alueita. Mineraalirakeet ovat epätasaisesti jakautuneina. Amadeuksen päämineraaleina on kalimaasälpä, kvartsi ja kordieriitti. Tarkempi mineraalikoostumus löytyy taulukosta 1.

*Taulukko 1. Amadeuksen mineraalikoostumus.*

Amadeus		
Kalimaasälpä	24,9	wt%
Kvartsi	21,2	wt%
Sillimaniitti	20	wt%
Kordieriitti	14,2	wt%
Biotiitti	10,2	wt%
Plagioklaasi	4,2	wt%
Granaatti	3,5	wt%
Ilmeniitti	1,7	wt%
Muut	0,03	wt%

#### **Balmoral Red, Taivassalon punainen graniitti**

Balmoral Red on Varsinais-Suomesta, Taivassalosta louhittu keski- ja karkearakeinen punainen graniitti. Balmoral Red on suuntautumaton ja tasarakenteinen magmakivi, jonka päämineraaleina on kalimaasälpä, plagioklaasi ja kvartsi. Yleisväriltään Balmoral on voimakkaan punainen. Balmoral Redin mineraalikoostumus löytyy taulukosta 2.

*Taulukko 2. Balmoral Redin mineraalikoostumus.*

Balmoral Red		
Kalimaasälpä	39,6	wt%
Kvartsi	30,1	wt%
Plagioklaasi	22,2	wt%
Biotiitti	5,2	wt%
Muskoviitti	1,3	wt%
Muut	1,8	wt%

### **Baltic Brown, Ylämaan ruskea graniitti**

Baltic Brown on Etelä-Karjalasta, Ylämaalta louhittu karkearakeinen, ruskea viborgiittigraniitti. Baltic Brown on suuntautumaton magmakivi ja sen ulkonäölle on ominaista suuret maasälpäkiteet, joita ympäröivät plagioklaasirenkaat. Balticin Brownin päämineraaleja ovat kalimaasälpä, plagioklaasi ja kvartsi. Tarkemmat tiedot Baltic Brownin mineraalikoostumuksesta löytyvät taulukosta 5.3.

*Taulukko 3. Baltic Brownin mineraalikoostumus.*

Baltic Brown		
Kalimaasälpä	36,8	wt%
Kvartsi	27,8	wt%
Plagioklaasi	25,9	wt%
Amfiboli	6,2	wt%
Biotiitti	2,1	wt%
Muut	1,3	wt%

### **Kuru Grey, Kurun harmaa graniitti**

Kuru Grey on Pirkanmaalta, Kurun Niemenkylästä louhittu hienorakeinen harmaa graniitti. Kuru Grey on suuntautumaton ja tasarakeinen magmakivi. Kuru Greyn päämineraalit ovat plagioklaasi, kvartsi ja kalimaasälpä. Kuru Greyn mineraalikoostumus on taulukossa 4.

*Taulukko 4. Kuru Greyn mineraalikoostumus.*

Kuru Grey		
Plagioklaasi	32,6	wt%
Kvartsi	31	wt%
Kalimaasälpä	30,8	wt%
Biotiitti	3,5	wt%
Muskoviitti	1,2	wt%
Muut	0,9	wt%

### **Lappia Blue, Kangasniemen siniharmaa graniitti**

Lappia Blue on Etelä-Savosta, Hirvensalmelta louhittu keski- ja karkearakeinen, suuntautumaton tummanharmaa graniitti. Lappia Bluen ulkonäölle ominaista ovat 10-20 mm kokoiset maasälpäkiteet. Lappia Bluen päämineraalit on kalimaasälpä, plagioklaasi, kvartsi ja biotiitti. Tarkka kuvaus Lappia Bluen mineraalikoostumuksesta löytyy taulukosta 5.

*Taulukko 5. Lappia Bluen mineraalikoostumus.*

Lappia Blue		
Kalimaasälpä	46	wt%
Plagioklaasi	21	wt%
Kvartsi	16	wt%
Biotiitti	14	wt%
Muut	3	wt%

**Lieto Red, Liedon punainen migmatiitti**

Lieto Red on Varsinais-Suomesta, Liedosta louhittu keskirakeinen punainen migmatiitti. Lieto Red on suuntautunut metamorfinen kivi, jonka pääväri on punainen. Lieto Redin päämineraalit ovat kalimaasälpä, kvartsi ja plagioklaasi. Tarkemmat tiedot Lieto Redin mineraalikoostumuksesta löytyvät taulukosta 6.

*Taulukko 6. Lieto Redin mineraalikoostumus.*

Lieto Red		
Kalimaasälpä	40,8	wt%
Kvartsi	33,8	wt%
Plagioklaasi	22	wt%
Biotiitti	3,2	wt%
Muut	0,1	wt%

**Nilsin Kvartsiitti**

Nilsin kvartsiitti on Pohjois-Savosta, Nilsistä louhittu hienorakeinen muskoviittikvartsiitti. Nilsin Kvartsiitti on metamorfinen kivi, jonka rakenne on hienorakeinen, suuntautunut ja liuskeinen. Nilsin kvartsiitin päämineraaleja ovat kvartsi ja muskoviitti. Mineraalikoostumuksen tarkempi kuvaus löytyy taulukosta 7.

*Taulukko 7. Nilsin kvartsiitin mineraalikoostumus.*

Nilsin Kvartsiitti		
Kvartsi	74,9	wt%
Muskoviitti	24,1	wt%
Muut	1,0	wt%

## 5.2 Mittaukset

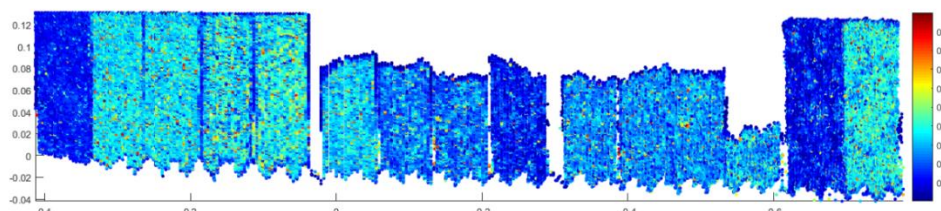
Mittaukset suoritettiin Paikkatietokeskuksen tiloissa Masalassa. Mittaukset tehtiin kahdessa erässä noin 4.8 metrin etäisyydeltä näytteistä. Laitteistona käytettiin kappaleessa 2.2.2 esiteltyä Paikkatietokeskuksen kehittämään hyperspektristä superjatkumo-lidaria. Näytteet oli aseteltu kuvan 20 mukaisesti betoniharkkojen kannattelevalle lankulle. Osa kivinäytteistä oli käännetty kiillotettu puoli mittalaitteistoon päin. Kuvan oikeassa laidassa näkyvä valkoinen paneeli on intensiteetin reflektanssiksi muuttamiseen käytetty Spectralon-paneeli.



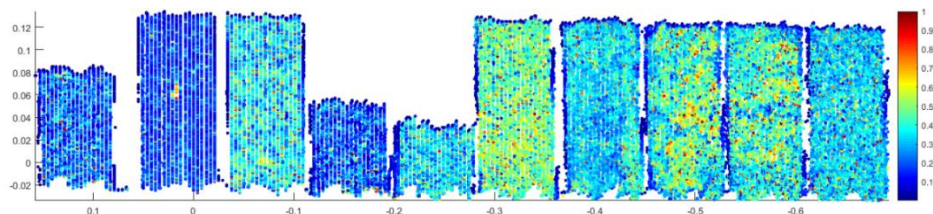
Kuva 20. Mittaus 1.

## 5.3 Datan käsittely

Datan esikäsittely tehtiin Paikkatietokeskuksen toimesta. Esikäsittely sisälsi aaltomuodon ylityksen ja jälkivärähtelyn poistamisen, pulssin intensiteetin normalisoinnin, Gaussisen pulssin sovituksen ja intensiteetin muuntamisen reflektanssiksi. Intensiteetin muuntaminen reflektanssiksi tehtiin 99 prosenttia heijastavan Spectralon-paneelin avulla. Lopullinen data sisälsi x, y ja z koordinaatit sekä reflektanssiarvot kahdeksalle kanavalle (564.3, 610.8, 659.9, 720.3, 764.8, 818.0, 878.6 ja 979.2 nm). Kanavien puoliarvoveveys (FWHM) oli noin 30 nm. Alkuperäinen data (mittaukset 1 ja 2) sisälsi 71410 ja 46544 pistettä. Datan esikäsittely oli laskennallisesti hidas prosessi, joka vei näiden mittausten osalta noin neljä päivää. Kuvassa 21 on nähtävissä ensimmäisen mittauksen pistepilvestä Matlabilla tehty kuva. Kuvan reflektanssiarvot ovat kanavalta 1 (564.3 nm). Kuvassa tummimpana näkyvät kivet ovat käännetty kiillotettu puoli mittalaitteistoon päin. Kuvassa 22 on toisen mittauksen pistepilvestä tehty kuva samalta aallonpituuskanavalta. Erityisesti vasemmalta kuudentena oleva kivi (Nilsin kvartsiitti) erottuu korkean reflektanssinsa johdosta.

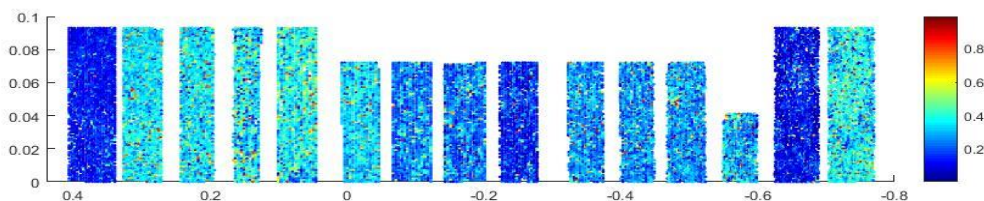


Kuva 21. Mittaus 1, kanava 1 (564.3 nm).



Kuva 22. Mittaus 2, kanava 1 (564.3 nm).

Kuten kuvista 21 ja 22 voidaan huomata, kivien ylä- ja alaosissa sekä sivuilla on pisteitä, jotka poikkeavat reflektanssiltaan selkeästi kiven keskiosista. Näiden osittaisista osumista johtuvien pisteiden poisto tehtiin leikkaamalla jokaisen kiven reunat pois ennen harjoitussarjojen ja tuntemattomien havaintojen keräämistä. Myös negatiiviset ja yli yhden arvoiset reflektanssiarvot poistettiin pistepilvestä. Tämän lisäksi päällekkäiset havainnot (sama x ja y koordinaatti) poistettiin. Ensimmäisen mittauksen käsitelty pistepilvi on nähtävissä kuvassa 23. Käsitelyn jälkeen pisteitä oli jäljellä 20792 kappaletta.



Kuva 23. Mittaus 1 käsitelty, kanava 1 (564.3 nm)

## 5.4 Harjoitussarjat ja tuntemattomat havainnot

Analyysejä ja luokittelua varten jokaisesta kivilajista valittiin yksi edustava näyte. Näyte jaettiin vertikaalisesti kahteen osaan, joista toista käytettiin harjoitussarjojen keräämiseen ja toista tuntemattomiin havaintoihin. Taulukkoon 8 on listattu valitut näytteet, niiden kuvaus ja myöhemmin tässä työssä luokittelussa käytettävä tunniste. Tarkempi kuvaus näytteistä löytyy kappaleesta 5.1.

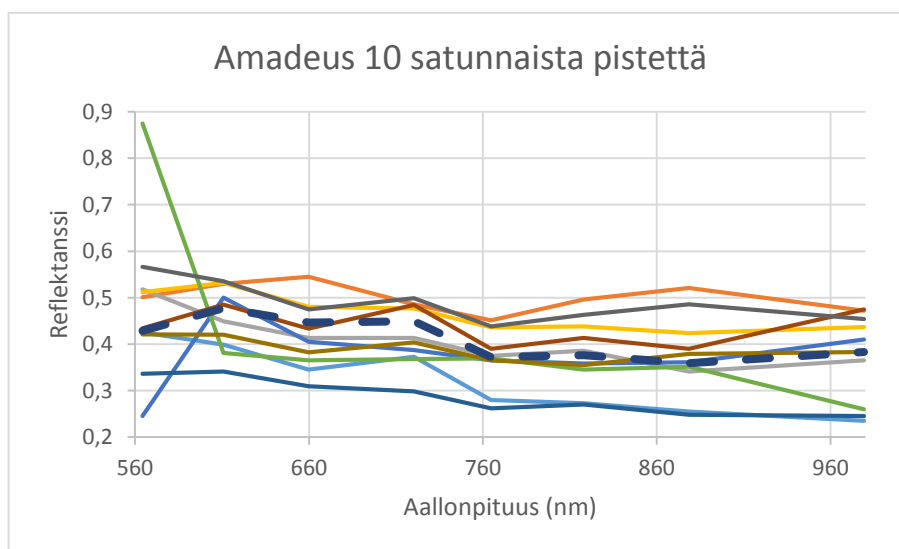


Taulukko 8. Analyyseissä käytetyt kiviläytteenä.

Näyte	Kuvaus	Luokittelu tunniste
Amadeus 3B	Granaattikordieriiditgneissi	Amadeus
Balmoral Red 3A	Keskirakeinen punainen graniitti	Balmoral
Baltic Brown 12A	Karkearakeinen ruskea viborgiittigraniitti	Baltic
Kuru Grey 8A	Hienorakeinen harmaa graniitti	Kuru
Lappia Blue 12A	Keskirakeinen siniharmaa graniitti	Lappia
Lieto Red 9A	Punainen migmatiitti	Lieto
Nilsian Kvartsiitti 15A	Muskoviittikvartsiitti	Nilsia

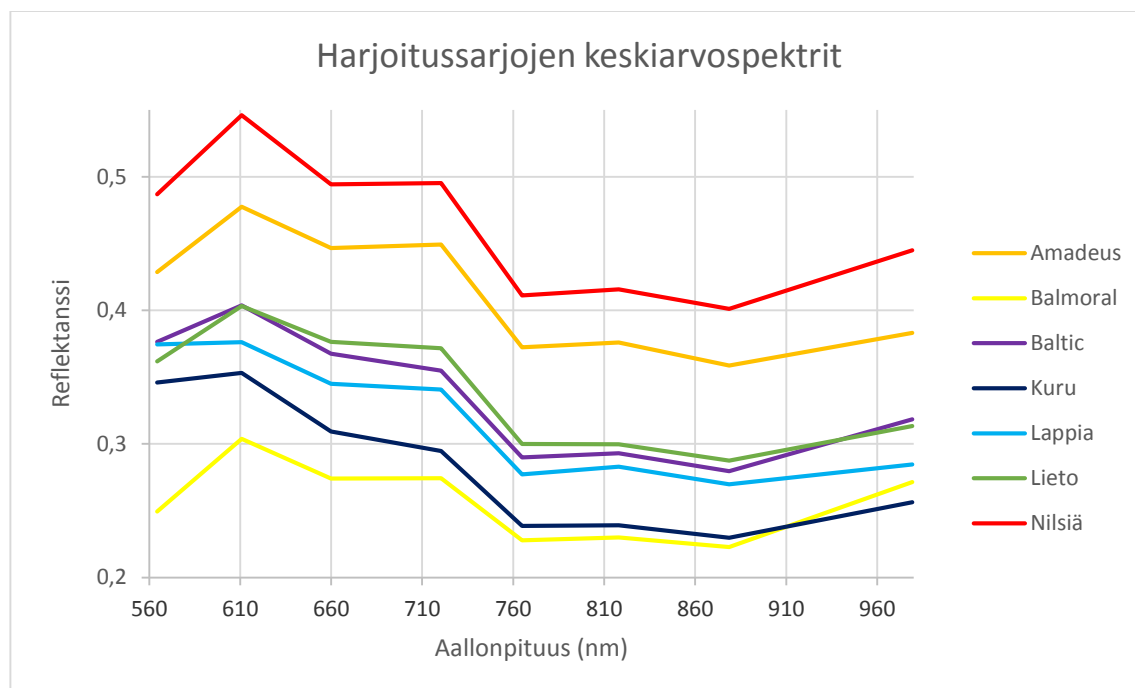
Harjoitussarjoja tehtiin 100 kappaletta kustakin kiviläytteenä (yhteensä 700 kappaletta). Jokainen harjoitusspektri koostui 200:sta satunnaisesti valitusta mittauspisteestä, jotka sijaitsivat kyseisen näytteen harjoitusalueella. Syynä keskiarvospektrien käyttöön oli yksittäisten mittauspisteiden spektrien reflektanssin ja käyrämuotojen hajonta, joka teki niistä käytännössä mahdottomia tunnistaa. Kuva 24 havainnollistaa edellä mainittua ongelmaa kymmenen satunnaisesti valitun Amadeuksen mittauspisteen spektrien kautta. Kuvassa 24 näkyvä katkoviiva kuvaa kaikkien Amadeus-luokan harjoitussarjojen keskiarvoa.

Luokitteluvaiheessa ilmeni, että harjoitus- ja tuntemattomiin spektreihin käytettyjen mittauspisteiden määrä ja valintatapa oli merkittävässä roolissa saatujen tulosten kannalta. Liian vähäisistä mittauspisteistä kootut harjoitusspektrit kasvattivat luokkien sisäisen varianssin niin suureksi, että niiden luokittelu kävi epävarmaksi. Kymmenen mittauspisteen keskiarvoina koostettujen harjoitusspektrien luokan sisäinen varianssi oli noin 20 kertaa ja sadan mittauspisteen keskiarvoista koostettujen 2 kertaa suurempi, kuin 200 pisteen keskiarvoista koostettujen. Myös mittauspisteiden satunnainen valinta pienensi luokkien sisäisiä variansseja.



Kuva 24. Amadeus-luokan harjoitusalueelta satunnaisesti valitut 10 spektrikäyriä.

Kuvasta 25 löytyy kaikkien harjoitussarjaluokkien keskiarvospektrit. Erityistä huomiota herättää kaikkien luokkien spektrikäyrien muodon samankaltaisuus. Erot eri luokkien välillä ovat selkeästi enemmän kokonaisreflektanssista riippuvaisia, kuin spektrikäyrän muodosta. Perinteisten luokittelijoiden yhteydessä käytetyt referenssispektrit tehtiin harjoitussarjojen keskiarvoista.



Kuva 25. Harjoitussarjojen keskiarvospektrit.

Tuntemattomat näytteet hankittiin samoista kiviläytteistä kuin harjoitussarjat, mutta käyttämällä kiven toista puolikasta. Kuten harjoitussarjat, myös tuntemattomat näytteet koostettiin 200:n satunnaisesti valitun mittauspisteen keskiarvosta. Tarkoituksena oli imitoida mittaustilannetta, jossa tuntemattomista kiviläytteistä mitattaisiin 200 pistettä ja ne luokiteltaisiin kyseisten pisteiden keskiarvojen perusteella ennalta määrättyihin luokkiin. Myös tuntemattomia havaintoja luotiin 100 kappaletta per kivilaji (yhteensä 700 kappaletta). Harjoitussarjojen ja tuntemattomien näytteiden sarjoissa oli kivilajista riippuen hieman eroavaisuuksia. Taulukossa 9 on listattuna harjoitussarjojen ja tuntemattomien näytteiden keskiarvojen eroavaisuudet eri aallonpituuksilla sekä niiden summa. Luvuista on havaittavissa, että luokat Amadeus ja Nilsia poikkeavat eniten keskiarvojen suhteen. Pienimmät eroavaisuudet ovat luokilla Baltic ja Lappia.

Taulukko 9. Harjoitussarjojen ja tuntemattomien havaintojen reflektanssien keskiarvojen erotukset eri aallonpituuksilla.

	564,3	610,8	659,9	720,3	764,8	818	878,6	979,2	Summa
Amadeus	0,011	0,013	0,007	0,005	0,006	0,009	0,013	0,019	0,083
Balmoral	0,014	0,007	0,006	0,003	0,005	0,003	0,002	0,001	0,041
Baltic	0,003	0,001	0,001	0,002	0,003	0,003	0,004	0,011	0,029
Kuru	0,006	0,010	0,008	0,008	0,006	0,007	0,007	0,006	0,057
Lappia	0,006	0,000	0,004	0,004	0,001	0,001	0,000	0,003	0,020
Lieto	0,000	0,002	0,005	0,008	0,007	0,007	0,004	0,006	0,039
Nilsia	0,010	0,000	0,018	0,023	0,016	0,014	0,018	0,009	0,107

## 5.5 Luokittelun tarkkuuden arviointi

Luokittelun tarkkuuden arviointi on olennainen osa saatujen tulosten tulkintaa. Luokittelun tarkkuus ei ole pelkästään riippuvainen valitusta algoritmista vaan myös mitatun datan laadulla on suuri merkitys luokittelun onnistumiseen. Luokittelun arvioinnilla voidaan kuitenkin verrata algoritmien keskinäistä hyvyttä kyseisen aineiston osalta. Tässä työssä käytetyt luokittelun tarkkuuden arviointimenetelmät ovat virhematriisi, kokonaistarkkuus ja kappa-arvo. Tässä kappaleessa esitellään myös tuottajan ja käyttäjän tarkkuudet, joiden avulla voidaan laskea oikeaan luokkaan luokittumisen ja luokittelun oikeellisuuden todennäköisyyksiä. Seuraavaksi esitellään edellä mainitut menetelmät.

Virhematriisi (kuva 26) on taulukko, josta ilmenee mihin referenssiluokkaan tuntemattomat näytteet ovat luokittuneet. Oikeisiin luokkiin luokittuneet näytteet ovat taulukon diagonaalilla. Virhematriisi on erittäin hyvä työkalu luokittelun arviointiin, koska sen avulla on mahdollista tarkastella oikein luokittuneiden näytteiden lisäksi, mihin luokkaan väärin luokittuneet näytteet ovat luokittuneet. Kuvasta 26 voidaan nähdä miten luokittelu tulokset sijoittuvat virhematriisissa. Kuvan 26 esimerkissä luokkaan 1 kuuluvista näytteistä 54 kappaletta on luokittunut oikein, 12 kappaletta on luokittunut luokkaan 2, 1 kappaletta on luokittunut luokkaan yksi ja 7 kappaletta on luokittunut luokkaan 4. Luokkaan 1 kuuluvia näytteitä on yhteensä 74 kappaletta. Luokkaan 1 on luokittunut yhteensä 64 kappaletta.

Virhematriisista voidaan laskea myös niin sanotut tuottajan ja käyttäjän tarkkuudet. Tuottajan tarkkuus kertoo, kuinka suurella todennäköisyydellä johonkin luokkaan kuuluva näyte on luokittunut oikeaan luokkaan. Esimerkki virhematriisissa (kuva 26) tuottajan tarkkuus luokalle 1 on  $54/74 = 0.73$ . Käyttäjän tarkkuus kertoo, kuinka suurella todennäköisyydellä johonkin luokkaan luokiteltu näyte kuuluu siihen luokkaan. Esimerkki virhematriisissa käyttäjän tarkkuus luokalle 1 on  $54/64 = 0.84$ .

Kertoo mihin luokkaan näyte kuuluu						
	Luokka 1	Luokka 2	Luokka 3	Luokka 4	Yhteensä	
Kertoo mihin luokkaan näyte on luokitellut	Luokka 1	54	2	8	0	64
	Luokka 2	12	38	1	1	52
	Luokka 3	1	5	45	0	51
	Luokka 4	7	1	3	62	73
	Yhteensä	74	46	57	63	240

Kuva 26. Esimerkki virhematriisista.

Kokonaistarkkuus on virhematriisin diagonaalilla olevien arvojen summa, jaettuna näytteiden kokonaismäärällä. Esimerkissä kokonaistarkkuus on  $(54+38+45+62)/240 = 0.83$ .

Kappa-arvo kertoo, kuinka paljon parempi luokittelun tulos on verrattuna satunnaiseen luokitteluun. Kappa saa arvoja väliltä 0..1. Kappa-arvot välillä 0.6-0.8 kertovat kohtuullisesta tarkkuudesta ja arvot jotka ovat yli 0.8 kertovat hyvästä tarkkuudesta. Kappa-arvo voidaan laskea kaavalla:

$$\hat{K} = \frac{n \sum_{i=1}^k n_{ii} - \sum_{i=1}^k n_{i+} n_{+i}}{n^2 - \sum_{i=1}^k n_{i+} n_{+i}}, \quad (20)$$

missä  $\sum n_{ii}$  on virhematriisin diagonaalilla olevien arvojen summa,  $n_{i+}$  on rivin  $i$  summa ja  $n_{+i}$  on sarakkeen  $i$  summa. (Wang, Weng 2013)

## 6 Tulokset

Seuraavaksi esitellään tehtyjen analyysien tulokset. Eri analyysimenetelmät esitellään erikseen ja niiden vertailu suoritetaan kappaleessa 7.

### 6.1 Spektrikulman mukainen luokittelu

Spektrikulman mukainen luokittelu tehtiin käyttämällä referenssinä harjoitussarjojen keskiarvospektrejä. Keskiarvospektrit laskettiin jokaiselle luokalle sadasta referenssispektristä. Jokaisen keskiarvospekttrin ja tuntemattoman spektrin (100 kappaletta per luokka) välinen kulma laskettiin ja tuntematon luokiteltiin pienimmän spektrikulman omaavaan luokkaan. Spektrikulman luokittelu tehdään yleensä määräämällä jokin raja-arvo spektrikulmalle, jonka alle jäävät arvot luokitellaan luokkaan kuuluviksi. Sopivan raja-arvon määrittäminen kyseiselle aineistolle ei kuitenkaan onnistunut, koska yhtenäisen raja-arvon määrittäminen kaikille luokille jätti luokittelematta tai luokitteli useaan ryhmään monet näytteet. Taulukosta 10 on nähtävissä spektrikulmien minimi, maksimi ja keskiarvot eri tuntemattomien ja referenssien välillä. Huomionarvoista lasketuissa kulmissa on se, että ne ovat erittäin pieniä. Yleisesti käytetty raja-arvo on 5 astetta, mikä olisi tässä tapauksessa luokittanut miltei kaikki näytteet jokaiseen luokkaan kuuluvaksi.

*Taulukko 10. Spektrikulmien minimi-, maksimi- ja keskiarvot.*

		Amadeus	Balmoral	Baltic	Kuru	Lappia	Lieto	Nilsä
Amadeus T	Min	0,22	2,17	1,61	3,76	2,11	1,24	0,35
	Max	1,44	3,45	3,30	5,68	4,08	2,24	1,61
	Keskiarvo	0,81	2,70	2,37	4,70	3,11	1,59	0,83
Balmoral T	Min	1,85	0,33	1,55	3,66	2,81	2,09	1,08
	Max	3,41	3,43	4,42	6,81	5,85	3,89	3,07
	Keskiarvo	2,53	1,10	3,09	5,45	4,45	2,83	1,86
Baltic T	Min	2,01	3,37	0,30	1,15	0,94	0,92	1,67
	Max	3,71	5,59	1,80	3,58	2,72	3,01	3,37
	Keskiarvo	2,82	4,60	0,95	2,08	1,42	1,88	2,48
Kuru T	Min	3,61	5,13	1,59	0,21	1,56	2,55	3,19
	Max	5,84	7,54	3,74	1,51	2,97	4,98	5,46
	Keskiarvo	4,57	6,18	2,43	0,64	2,15	3,63	4,14
Lappia T	Min	1,71	4,33	1,15	1,44	0,13	1,31	1,84
	Max	4,48	6,65	2,95	3,27	1,57	3,89	4,31
	Keskiarvo	2,86	5,17	1,71	2,29	0,62	2,29	2,74
Lieto T	Min	1,20	3,01	0,65	1,83	1,03	0,27	1,16
	Max	2,95	4,92	2,16	4,28	3,13	2,01	2,65
	Keskiarvo	1,75	3,79	1,36	3,34	2,13	0,73	1,68
Nilsä T	Min	0,49	2,10	1,98	4,22	2,41	1,27	0,54
	Max	1,09	3,39	2,98	5,40	3,77	1,92	1,39
	Keskiarvo	0,78	2,71	2,47	4,81	3,16	1,60	0,96

Taulukosta 6.2 nähdään spektrikulman avulla tehty luokittelu, jossa tuntematon havainto on luokiteltu pienimmän spektrikulman arvon saavaan luokkaan. Spektrikulman avulla näytteet



## 6.2 Etäisyysluokittelu

Etäisyysluokittelu tehtiin kuten spektrikulman mukainen luokittelu, käyttämällä referensseinä keskiarvospektrejä ja luokittelemalla tuntematon näyte pienimmän etäisyyden saavaan luokkaan. Etäisyys laskettiin referenssispektrin ja tuntemattoman spektrin euklidisena etäisyytenä. Taulukossa 12 on etäisyysluokittelun minimi-, maksimi- ja keskiarvot. Arvoista on havaittavissa, että referenssispektrin ja tuntemattomien havaintojen välisten etäisyyksien keskiarvot ovat erittäin pieniä.

*Taulukko 12. Etäisyysluokittelun minimi-, maksimi- ja keskiarvot*

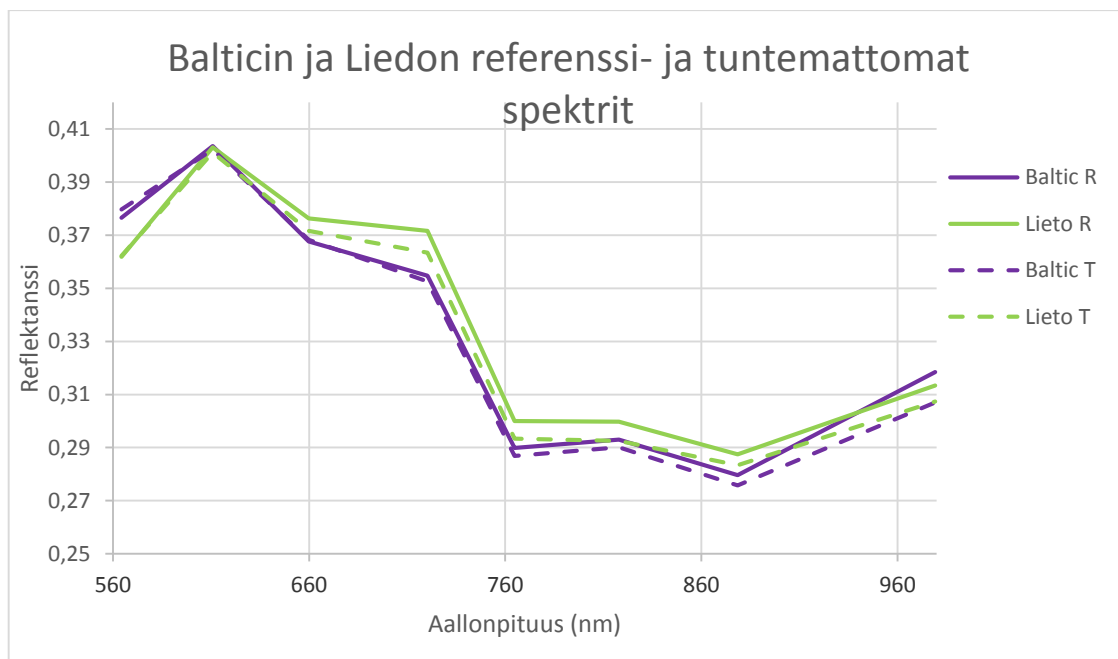
		Amadeus	Balmoral	Baltic	Kuru	Lappia	Lieto	Nilsia
Amadeus T	Min	0,01	0,43	0,20	0,35	0,25	0,19	0,08
	Max	0,07	0,51	0,28	0,43	0,33	0,27	0,16
	Keskiarvo	0,04	0,47	0,25	0,40	0,29	0,23	0,12
Balmoral T	Min	0,40	0,01	0,19	0,07	0,15	0,20	0,54
	Max	0,46	0,05	0,25	0,13	0,21	0,26	0,60
	Keskiarvo	0,43	0,02	0,22	0,10	0,18	0,23	0,57
Baltic T	Min	0,20	0,21	0,01	0,12	0,02	0,02	0,34
	Max	0,25	0,26	0,04	0,16	0,07	0,05	0,39
	Keskiarvo	0,23	0,23	0,02	0,14	0,05	0,04	0,37
Kuru T	Min	0,32	0,11	0,11	0,01	0,07	0,13	0,46
	Max	0,36	0,16	0,15	0,04	0,10	0,16	0,51
	Keskiarvo	0,35	0,13	0,13	0,02	0,08	0,15	0,49
Lappia T	Min	0,24	0,15	0,04	0,06	0,00	0,05	0,39
	Max	0,31	0,22	0,10	0,12	0,05	0,11	0,46
	Keskiarvo	0,27	0,19	0,06	0,10	0,02	0,07	0,41
Lieto T	Min	0,20	0,21	0,01	0,13	0,04	0,01	0,34
	Max	0,24	0,25	0,04	0,17	0,07	0,03	0,38
	Keskiarvo	0,22	0,23	0,02	0,15	0,06	0,02	0,36
Nilsia T	Min	0,16	0,60	0,38	0,53	0,43	0,37	0,03
	Max	0,21	0,65	0,42	0,57	0,47	0,41	0,07
	Keskiarvo	0,18	0,62	0,40	0,55	0,45	0,39	0,04

Taulukosta 13 nähdään etäisyysluokittelun tulokset. Luokittelun kokonaistarkkuus on 96.6 prosenttia, jota voidaan pitää erittäin hyvänä tuloksena. Amadeus, Balmoral, Kuru, Lappia ja Nilsia luokittelevat täydellisesti. Baltic luokituu 98 prosenttisesti ja Lieto 78 prosenttisesti oikein. Balticin kaksi väärää luokitusta menevät Lappia-luokkaan ja Liedon 22 väärää luokitusta menevät kaikki Baltic-luokkaan. Kuvasta 27 voidaan nähdä Balticin ja Liedon referenssispektrien sekä Balticin ja Liedon tuntemattomien havaintojen keskiarvospektrit.

*Taulukko 13. Etäisyysluokittelun tulokset.*

	Amadeus	Balmoral	Baltic	Kuru	Lappia	Lieto	Nilsia	Yhteensä
Amadeus	100	0	0	0	0	0	0	100
Balmoral	0	100	0	0	0	0	0	100
Baltic	0	0	98	0	0	22	0	120
Kuru	0	0	0	100	0	0	0	100
Lappia	0	0	2	0	100	0	0	102
Lieto	0	0	0	0	0	78	0	78
Nilsia	0	0	0	0	0	0	100	100
Yhteensä	100	100	100	100	100	100	100	700

Kuvan 27 perusteella voidaan helposti ymmärtää, miksi etäisyysluokittelija luokittelee osan tuntemattomista Lieto-luokkaan kuuluvista havainnoista Baltic-luokkaan ja miksi etäisyysluokittelija ei luokittele Baltic-näytteitä Lieto-luokkaan. Balticin tuntemattomien havaintojen keskiarvo on selkeästi lähempänä Balticin referenssispekttriä kuin Liedon referenssispekttriä, kun taas Liedon tuntemattomien havaintojen keskiarvo on Balticin ja Liedon referenssispektrien välissä. Kuvasta 27 on myös havaittavissa miksi spektrikulman avulla tehty luokittelu ei sekoittanut Balticia ja Lietoa yhtä pahasti kuin etäisyysluokittelu. Peräkkäisten aallonpituuksien välinen kulma on Balticin ja Liedon välillä kohtuullisen selkeästi toisistaan poikkeava.



*Kuva 27. Balticin ja Liedon referenssi- (R) ja tuntemattomat (T) spektrit.*



### 6.3 Korrelaatioluokittelu

Myös korrelaatioluokittelu tehtiin käyttämällä referensseinä keskiarvospektrejä ja luokittelemalla tuntematon havainto suurimman korrelaation saavaan luokkaan. Taulukosta 14 nähdään Pearsonin korrelaatiokertoimet eri referenssispektrien ja tuntemattomien näytteiden välillä. Arvoista voidaan todeta, että eri luokkiin kuuluvat näytteet korreloivat erittäin voimakkaasti.

*Taulukko 14. Korrelaatioluokittelun minimi-, maksimi- ja keskiarvot*

		Amadeus	Balmoral	Baltic	Kuru	Lappia	Lieto	Nilsia
Amadeus T	Min	0,97	0,82	0,92	0,84	0,85	0,97	0,97
	Max	1,00	0,93	1,00	0,98	0,98	1,00	1,00
	Keskiarvo	0,99	0,89	0,97	0,92	0,93	0,99	0,99
Balmoral T	Min	0,83	0,85	0,80	0,68	0,66	0,83	0,88
	Max	0,95	1,00	0,99	0,96	0,93	0,96	0,98
	Keskiarvo	0,91	0,98	0,91	0,82	0,80	0,91	0,95
Baltic T	Min	0,90	0,71	0,98	0,93	0,93	0,93	0,92
	Max	0,99	0,90	1,00	1,00	0,99	1,00	0,99
	Keskiarvo	0,96	0,81	0,99	0,98	0,98	0,97	0,97
Kuru T	Min	0,86	0,66	0,95	0,99	0,98	0,89	0,89
	Max	0,96	0,81	0,99	1,00	0,99	0,97	0,97
	Keskiarvo	0,92	0,75	0,98	1,00	0,99	0,94	0,94
Lappia T	Min	0,86	0,62	0,94	0,96	0,98	0,89	0,87
	Max	0,98	0,78	0,99	0,99	1,00	0,98	0,97
	Keskiarvo	0,93	0,72	0,97	0,99	1,00	0,95	0,94
Lieto T	Min	0,96	0,78	0,96	0,90	0,91	0,97	0,97
	Max	1,00	0,90	1,00	0,99	0,99	1,00	1,00
	Keskiarvo	0,99	0,85	0,98	0,95	0,96	1,00	0,99
Nilsia T	Min	0,98	0,83	0,93	0,86	0,88	0,98	0,97
	Max	1,00	0,94	0,99	0,95	0,97	1,00	1,00
	Keskiarvo	0,99	0,89	0,97	0,91	0,93	0,99	0,99

Taulukosta 15 nähdään suurimman korrelaatiokertoimen avulla tehty luokittelu. Luokittelun kokonaistarkkuus on 69.7 prosenttia, mikä on luokittelutuloksista heikoin. Korrelaatiokertoimen avulla Kuru ja Lappia luokittuivat melkein täydellisesti oikein. Balmoral ja Lieto luokittuivat yli 80 prosenttisesti ja Baltic 75 prosenttisesti. Amadeus ja Nilsia erottuivat selvästi heikoiten, niiden luokittelutarkkuuden ollessa alle 30 prosenttia. Amadeuksen tuntemattomien näytteiden sarjasta 21 luokittui Liedoksi ja 46 Nilsiäksi. Nilsian tuntemattomista 38 luokittui Amadeukseksi ja 42 luokittui Liedoksi.

*Taulukko 15. Korrelaatioluokittelun tulokset.*

	Amadeus	Balmoral	Baltic	Kuru	Lappia	Lieto	Nilsia	Yhteensä
Amadeus	28	0	0	0	0	1	38	67
Balmoral	0	82	0	0	0	0	0	82
Baltic	5	4	75	3	0	9	0	96
Kuru	0	0	18	97	1	0	0	116
Lappia	0	0	2	0	98	1	0	101
Lieto	21	0	4	0	1	88	42	156
Nilsia	46	14	1	0	0	1	20	82
Yhteensä	100	100	100	100	100	100	100	700

## 6.4 Kanoninen erotteluanalyysi

Kanonisessa erotteluanalyysissä käytettiin harjoitussarjoina 700 kappaletta referenssispektrejä (100 kappaletta per luokka). Harjoitussarjojen avulla luotu malli erotteli referenssispektrit täydellisesti oikeisiin luokkiin. Mallille tehtiin myös ristiin validointi, jossa jokainen referenssispektri poistettiin yksitellen harjoitussarjoista ja luokiteltiin sen jälkeen johonkin luokkaan kuuluvaksi. Ristiin validoinnin virheprosentti oli nolla, joten mallin ennustavuus oli erittäin hyvä.

Taulukosta 16 on nähtävissä kanonisen erotteluanalyysin tuottamien muuttujien ominaisarvot ja niiden osuudet kokonaisvarianssista. Ensimmäisen muuttujan ominaisarvo (305) ja osuus varianssista (88,8 %) on erittäin suuri. Toinen muuttuja selittää noin 8 prosenttia ja kolmas muuttuja noin 3 prosenttia varianssista. Muiden muuttujien merkitys mallille on hyvin pieni.

*Taulukko 16. Kanonisen erotteluanalyysin ominaisarvot ja kumulatiiviset osuudet.*

Muuttuja	Ominaisarvo	Prosenttia varianssista	Kumulatiivinen	Kanoninen Korrelaatio
1	305,46041	87.99 %	87.99 %	0,99837
2	28,13643	8.10 %	96.09 %	0,98269
3	9,1623	2.64 %	98.73 %	0,94952
4	3,2649	0.94 %	99.67 %	0,87494
5	0,87422	0.25 %	99.92 %	0,68297
6	0,26646	0.08 %	100.00 %	0,45869

Taulukossa 17 on Wilksin lambda-arvot ja khiin neliö -testin tulokset. Wilksin lambda on testisuure, joka kertoo kuinka hyvin erottelufunktiot pystyvät erottelemaan ryhmät toisistaan. Testisuure lasketaan ensin kaikilla erottelufunktiolla, jonka jälkeen testi ajetaan uudelleen jättäen aina yksi erottelufunktio pois testistä. Wilksin lambda saa arvoja väliltä 0-1. Lähellä nollaa olevat arvot kertovat hyvästä erottelukyvystä, kun taas ykkösen lähellä olevat arvot kertovat huonosta erottelukyvystä. Taulukosta 17 on havaittavissa, että Wilksin lambda-arvot ja saadut p-arvot (0.05 merkitsevyystasolla) tukevat oletusta kuuden erottelufunktion käytöstä, joten kaikki kuusi erottelufunktiota on otettu mukaan erotteluanalyysiin.

Taulukko 17. Wilksin lambda ja khiin neliö.

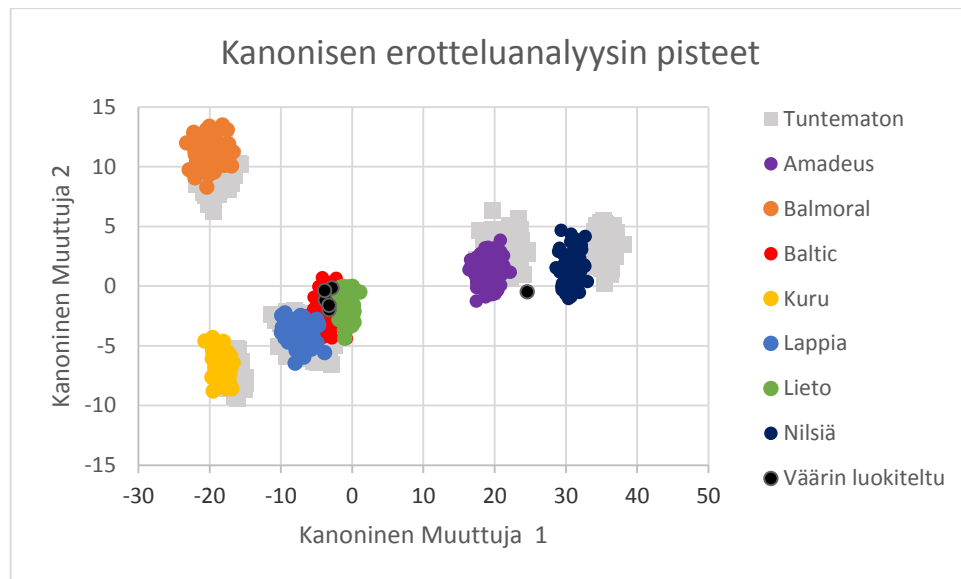
Erottelufunktiot	Wilksin lambda	Khiin neliö	Vapausasteet	p-arvo
1-6	1,09E-06	9494,70921	48	0
2-6	3,34E-04	5535,81047	35	0
3-6	0,00972	3204,07982	24	0
4-6	0,09878	1600,70903	15	0
5-6	0,4213	597,7443	8	7,15E-124
6-6	0,7896	163,34866	3	3,47E-35

Mallin luomisen ja ristiin validoinnin jälkeen 700 kappaletta tuntemattomia heijastusspektrejä (100 per luokka) luokiteltiin mallin avulla. Taulukosta 18 ovat nähtävissä luokittelutulokset. Luokkiin Balmoral, Baltic, Kuru, Lappia ja Nilsia kuuluvat tuntemattomat luokittuivat 100 prosenttisesti oikein. Luokkaan Amadeus kuuluvista tuntemattomista 99 luokittui oikein ja yksi meni luokkaan Nilsia. Luokkaan Lieto kuuluvista tuntemattomista spektreistä 95 meni oikeaan luokkaan ja 5 luokkaan Baltic. Mallin kokonaisluokittelutarkkuus oli yli 99 prosenttia, jota voidaan pitää erittäin hyvänä.

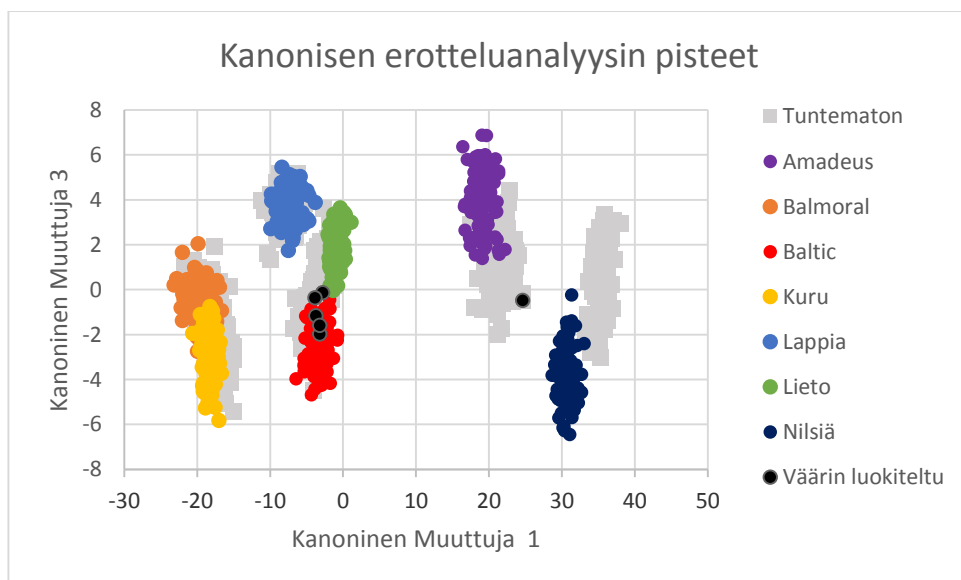
Taulukko 18. Tuntemattomien näytteiden luokittuminen kanonisessa erotteluanalyysissä.

	Amadeus	Balmoral	Baltic	Kuru	Lappia	Lieto	Nilsia	Yhteensä
Amadeus	99	0	0	0	0		0	99
Balmoral	0	100	0	0	0	0	0	100
Baltic	0	0	100	0	0	5	0	105
Kuru	0	0	0	100	0	0	0	100
Lappia	0	0	0	0	100	0	0	100
Lieto	0	0	0	0	0	95	0	95
Nilsia	1	0	0	0	0	0	100	101
Yhteensä	100	100	100	100	100	100	100	700

Kuvasta 28 nähdään kanonisen erotteluanalyysin tuottamien kanonisten muuttujien 1 ja 2 välinen projektio tasolle. Kuvasta on havaittavissa, että Balmoral, Kuru, Amadeus ja Nilsia ryhmittyvät selkeästi omiksi ryppäiksi. Luokat Lappia, Baltic ja Lieto asettuvat muuttujien 1 ja 2 virittämällä tasolla hyvin lähekkäin. Kuvassa vaaleanharmaina neliöinä esitetyt tuntemattomat näytteet asettuvat kohtuullisen hyvin harjoitussarjojen avulla luotujen luokkien alle. Kuvassa 29 on kanonisten muuttujien 1 ja 3 välinen projektio tasolle. Kuvasta on havaittavissa, että kolmas kanoninen muuttuja erottelee luokat Lappia, Baltic ja Lieto toisistaan. Kuvissa 28 ja 29 on myös nähtävissä väärin luokiteltujen (musta ympyrä) spektrien sijoittuminen muuttujatasoille. Erityisesti luokkaan Baltic menneiden Lieto-näytteiden sijoittuminen kuvan 29 kuvaamalle tasolle selittää niiden virheellisen luokittelun.



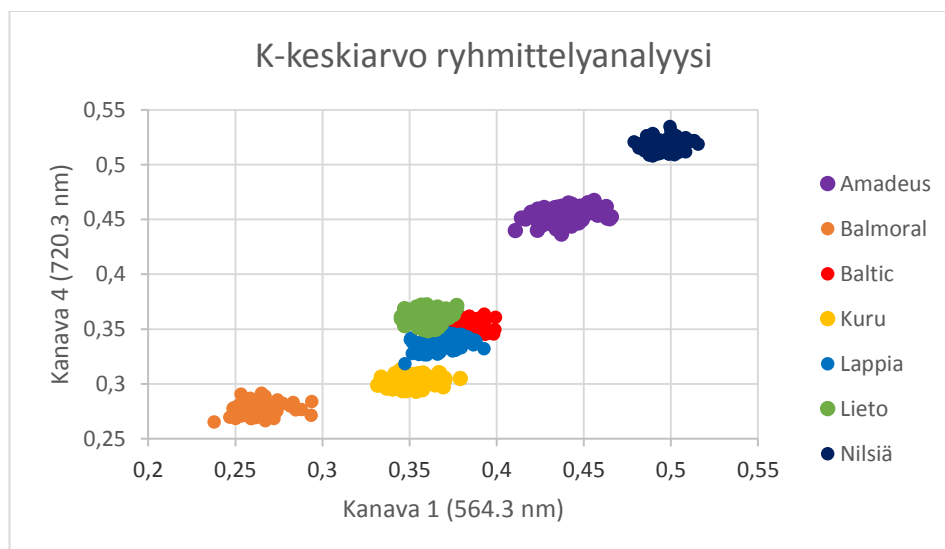
Kuva 28. Kanonisen erotteluanalyysin tulokset.



Kuva 29. Kanonisen erotteluanalyysin tulokset.

## 6.5 Ryhmittelyanalyysi

Ryhmittelyanalyysi tehtiin kahdessa osassa. Ensimmäisessä vaiheessa luotiin hierarkkisen ryhmittelyanalyysin avulla harjoitussarjoista seitsemän ryhmää, jotka kuvasivat seitsemää kiviluokkaa. Ryhmittelyssä käytettiin Wardin menetelmää, etäisyytenä euklidista etäisyyttä ja ryhmittymät etsittiin etäisyyksien summan perusteella. Hierarkkisella ryhmittelyanalyysillä saatujen ryhmien keskipisteiden avulla tehtiin k-keskiarvo -ryhmittelyanalyysi, jossa olivat mukana sekä harjoitussarjat että tuntemattomat havainnot. K-keskiarvo ryhmittelyanalyysin tulosten projektiomuuttujien kanava 1 ja kanava 4 virittämälle tasolle on nähtävissä kuvassa 30. Havainnot ovat selvästi ryhmittyneet omiin klustereihinsa. Kuten kanonisessa erotteluanalyysissä, myös ryhmittelyanalyysissä havainnot Baltic, Lappia ja Lieto ovat toistensa lähellä. Amadeus, Nilsia ja Balmoral erottuvat selkeimmin toisistaan. Luokkaan Kuru kuuluvat havainnot ryhmittyvät omaksi klusterikseen, mutta sijaitsevat lähellä Lappia-luokkaa.



Kuva 30. K-keskiarvo ryhmittelyanalyysi.

Taulukkoon 19 on koostettu ryhmien väliset etäisyydet. Etäisyyksistä on havaittavissa, että ryhmät Baltic, Lieto ja Lappia sijaitsevat erittäin lähellä toisiaan. Kauimpana toisistaan ovat ryhmät Balmoral ja Nilsia. Keskimääräinen etäisyys luokkien välillä on 0,255.

Taulukko 19. Ryhmien väliset etäisyydet.

	Amadeus	Nilsia	Balmoral	Kuru	Baltic	Lieto	Lappia
Amadeus	0	0,15186	0,45616	0,37481	0,25288	0,25023	0,30147
Nilsia	0,15186	0	0,6076	0,52507	0,40301	0,40122	0,45207
Balmoral	0,45616	0,6076	0	0,11457	0,21535	0,2116	0,16955
Kuru	0,37481	0,52507	0,11457	0	0,12235	0,1281	0,07666
Baltic	0,25288	0,40301	0,21535	0,12235	0	0,02518	0,05298
Lieto	0,25023	0,40122	0,2116	0,1281	0,02518	0	0,05855
Lappia	0,30147	0,45207	0,16955	0,07666	0,05298	0,05855	0

Talukossa 20 on K-keskiarvo ryhmittelyanalyysin tulokset. Tuloksissa on mukana ainoastaan tuntemattomien havaintojen spektrit. Amadeus, Balmoral, Kuru ja Nilsia ryhmittyyvät täydellisesti omiin ryhmiinsä. Balticin havainnoista 84 kappaletta ryhmittyy oikein ja 16 kappaletta luokituu Lieto-ryhmään. Lappian havainnoista 98 kappaletta menee oikeaan ryhmään ja kaksi kappaletta menee Baltic-ryhmään. Liedon havainnoista 94 kappaletta ryhmittyy oikein ja kuusi kappaletta luokituu Baltic-ryhmään. Kokonaistarkkuus ryhmittelyanalyysissä on 96,6 prosenttia, mitä voidaan pitää erittäin hyvänä.

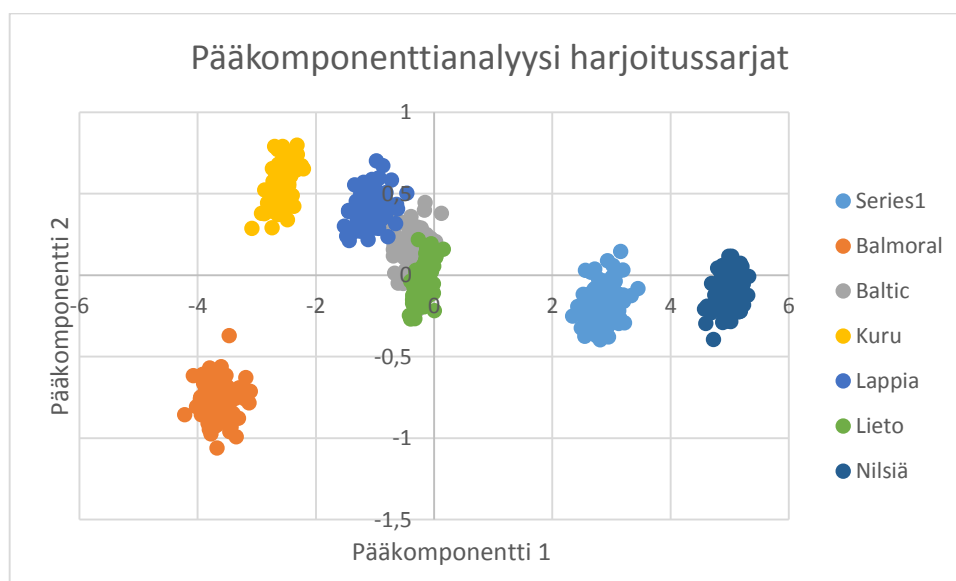
Taulukko 20.. K-keskiarvo ryhmittelyanalyysin tulokset.

	Amadeus	Balmoral	Baltic	Kuru	Lappia	Lieto	Nilsia	Yhteensä
Amadeus	100	0	0	0	0	0	0	100
Balmoral	0	100	0	0	0	0	0	100
Baltic	0	0	84	0	2	6	0	92
Kuru	0	0	0	100	0	0	0	100
Lappia	0	0	0	0	98	0	0	98
Lieto	0	0	16	0	0	94	0	110
Nilsia	0	0	0	0	0	0	100	100
Yhteensä	100	100	100	100	100	100	100	700

## 6.6 Pääkomponenttimuunnos

Pääkomponenttimuunnos on dimensioiden vähentämistekniikka, jonka tuottamat uudet muuttujat (pääkomponentit) ovat alkuperäisten muuttujien lineaarikombinaatioita. Pääkomponenttien suunta valitaan niin, että ne maksimoivat mittauspisteiden välisen varianssin. Tämän työn yhteydessä tehdyissä mittauksissa saatiin spektritietoa vain kahdeksalta kanavalta, joten dimensioiden vähentämiselle ei ollut suurta tarvetta. Harjoitussarjoille ja tuntemattomille näytteille tehtiin kuitenkin pääkomponenttimuunnosta ja sen toimivuutta testattiin ryhmittelyanalyysin avulla. Tarkoituksena oli selvittää, vaikuttaako dimensioiden vähentäminen kahdeksasta kolmeen merkittävästi ryhmittelyanalyysin tuloksiin.

Pääkomponenttimuunnos tehtiin ensin harjoitussarjoille (kuva 31), ja siitä saatuja kolmea pääkomponenttia käytettiin hierarkkiseen ryhmittelyanalyysiin. Ryhmittelyanalyysistä saadut ryhmien keskiarvot asetettiin k-keskiarvo ryhmittelyanalyysiin ryhmien keskipisteiksi. Seuraavaksi tehtiin pääkomponenttimuunnos tuntemattomille näytteille (kuva 32) ja siitä saadut kolme pääkomponenttia käytettiin k-keskiarvo ryhmittelyanalyysin muuttujina.



Kuva 31. Pääkomponenttianalyysi harjoitussarjoille



## 7 Tulosten tarkastelu ja pohdinta

Luokittelu suoritettiin kolmella perinteisellä luokittelumenetelmällä (spektrikulma, etäisyys ja korrelaatiokerroin), yhdellä erotteluanalyysillä ja yhdellä ryhmittelyanalyysillä. Dimensioiden vähentämisen vaikutuksia testattiin pääkomponenttimuunnoksen ja sen tuloksien perusteella tehdyllä ryhmittelyanalyysillä. Perinteisistä luokittelumenetelmistä yksinkertainen euklidiseen etäisyyteen perustuva luokittelu osoittautui kaikkein tarkimmaksi. K-keskiarvo ryhmittelyanalyysi oli yhtä tarkka kuin etäisyysluokittelija. Kaikkein parhaimmat luokittelutulokset saatiin kanonisella erotteluanalyysillä, jonka luokittelutarkkuus oli yli 99 prosenttia. Taulukossa 22 on listattuna kaikkien luokittelumenetelmien kokonaistarkkuudet, kappa-arvot ja heikoiten luokitteleva kiviluokka.

*Taulukko 22. Luokittelumenetelmät ja niiden tarkkuus.*

Luokittelumenetelmä	Tarkkuus	Kappa	Heikoin
Spektrikulma	74,3 %	0,70	Nilsia
Etäisyysluokittelu	96,6 %	0,96	Lieto
Korrelaatiokerroin	69,7 %	0,65	Nilsia
Kanoninen erotteluanalyysi	99,1 %	0,99	Lieto
K-keskiarvo ryhmittelyanalyysi	96,6 %	0,96	Baltic ja Lieto
Pääkomponenttianalyysi	95,9 %	0,95	Baltic

Luokittelun onnistuminen ei johdu pelkästään luokittelumenetelmästä, vaan myös harjoitusarjojen ja tuntemattomien näytteiden samankaltaisuus ja etäisyys muista luokista vaikuttaa merkittävästi tulokseen. Taulukkoon 23 on koottu tuntemattomien näytteiden luokittuminen eri menetelmillä. Keskiarvon perusteella heikoiten kaikista kivilajeista luokitteleva Nilsia, joka kuitenkin luokiteltiin kolmella eri menetelmällä täydellisesti. Käytännössä Lieto oli vaikein luokiteltava, koska sitä ei saatu millään menetelmällä luokiteltua täydellisesti. Helpoin luokiteltava oli Kuru, joka luokitteleva neljällä menetelmällä täydellisesti ja yhdellä 97 prosenttisesti.

*Taulukko 23. Tuntemattomien näytteiden luokittuminen eri menetelmillä.*

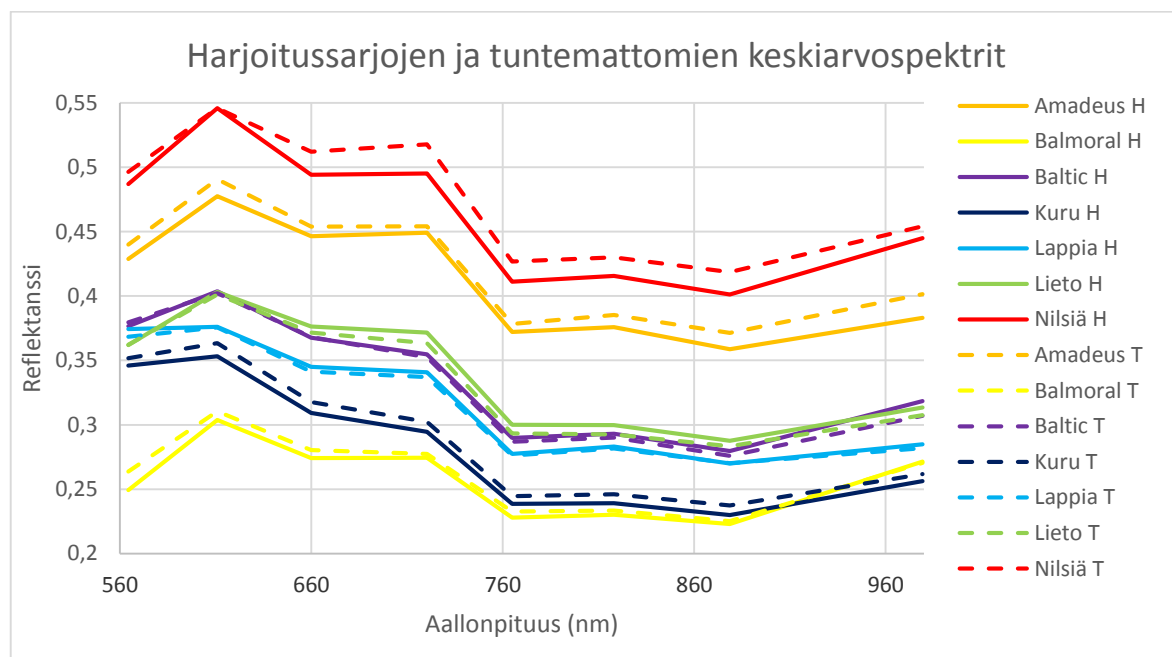
	Amadeus	Balmoral	Baltic	Kuru	Lappia	Lieto	Nilsia
Spektrikulma	53	82	77	100	99	87	22
Etäisyys	100	100	98	100	100	78	100
Korrelaatio	28	82	75	97	98	88	20
Kanoninen	99	100	100	100	100	95	100
K-keskiarvo	100	100	84	100	98	94	100
Keskiarvo	76,00 %	92,80 %	86,80 %	99,40 %	99,00 %	88,40 %	68,40 %

Tulosten tarkastelun yhteydessä havaittiin, että muuttujien lineaarista riippuvuutta mittaavat luokittelijat (spektrikulma ja korrelaatio) olivat herkkiä harjoitusarjojen ja tuntemattomien näytteiden välisille spektrikäyrän muodon eroavaisuuksille (kuva 33), kun taas muuttujien etäisyyteen perustuvat menetelmät, jotka luokittelivat lähimpään mahdolliseen ryhmään, olivat herkkiä ryhmien lähekkäisyydelle. Tästä voidaan päätellä, että spektrikulman tai korrelaation avulla tehtyä luokitusta voi selvästi parantaa luokkaa paremmin edustavilla harjoitusarjoilla. Tulosten tarkastelun yhteydessä spektrikulmat ja korrelaatio laskettiin myös harjoitusarjojen ja referenssispektrin (harjoitusarjojen keskiarvo) välille, jolloin spektrikulman luokittelutarkkuudeksi saatiin Amadeukselle 96 prosenttia ja Nilsialle 100 prosenttia.



Korrelaatioluokitusten tulokset Amadeuksen ja Nilsin harjoitussarjoille olivat 79 prosenttia ja 98 prosenttia. Tämä tukee olettamusta harjoitussarjojen valinnan tärkeydestä. Liitteen 1 kuvasta 1 nähdään, että Amadeuksen mineraalirakeiden spatiaalinen jakautuminen näytteen vertikaalisuunnassa kahtia jaetuille puoliskoille on erittäin heterogeenistä. Sen perusteella on helppo käsittää eroavaisuudet Amadeuksen harjoitussarjan ja tuntemattomien havaintojen välillä. Huomattavasti erikoisempia ovat koostumukseltaan hyvin homogeenisen Nilsin harjoitussarjojen ja tuntemattomien havaintojen välillä olevat eroavaisuudet, joille ei löydetty selkeää selitystä.

Ryhmien välisiin etäisyyksiin perustuvien menetelmien luokittelutuloksia pystyy lähinnä parantamaan poistamalla poikkeavia havaintoja, jotka ovat herkimpiä luokittumaan väärin. Poikkeavien havaintojen poistamisessa tulee kuitenkin olla varovainen, koska yksittäisten mittauspisteiden välillä on niin merkittäviä eroja, että tärkeää tietoa voidaan menettää, jos karsinta tehdään liian tiukoilla reunaehdoilla. Kuvasta 33 ovat nähtävissä sekä harjoitussarjojen että tuntemattomien havaintojen keskiarvospektrit. Kuvan perusteella on helppo ymmärtää, miksi ryhmien etäisyyksiin perustuvat menetelmät toimivat selkeästi paremmin Amadeuksen ja Nilsin kohdalla kuin Balticin ja Liedon kohdalla.



Kuva 33. Harjoitussarjojen ja tuntemattomien näytteiden keskiarvospektrit

Kuten kappaleessa 5.4 mainittiin, harjoitussarjojen ja tuntemattomien havaintojen keräämiseen käytetty menetelmä ja mittauspisteiden määrä vaikutti selkeästi luokittelun tuloksiin. Useiden mittauspisteiden keskiarvojen käyttäminen luokittelussa vähensi luokkien sisäistä varianssia ja poikkeavien havaintojen määrää. Satunnaisotannalla valitut mittauspisteet tuottivat selvästi paremman tuloksen kuin vierekkäisten pisteiden käyttö. Tämä korostuu erityisesti suurissa mineraalirakeita sisältävissä kivistä, koska viereisten pisteiden käytöllä saataan saada ylisuuri edustus yhdestä mineraaliluokasta.

Tehdyissä mittauksissa olivat rajoittavina tekijöinä kanavien määrä ja aallonpituusulottuvuus. Tyypillisesti hyperspektrikuviksi lasketaan vähintään kymmeniä kanavia sisältävät ja

yhtenäisen spektrin muodostavat kuvat. Tässä työssä käytetyt kahdeksan kanavaa riittivät seitsemän kivilajin erottamiseen toisistaan, mutta jos kivilajien määrää olisi kasvatettu kymmeniin tai satoihin, olisivat luokkien väliset etäisyydet luultavasti menneet hyvin pieniksi ja ristiin luokittumista olisi tapahtunut enemmän. Kuten tämän työn teoriaosuudessa mainittiin, mineraalien ja kivien karakteristiset ominaisuudet ovat aallonpituusalueella 400–2500 nm. Koska tehtyjen mittausten aallonpituusulottuvuus oli ainoastaan välillä 564–979 nm, jäi suuri osa karakteristisista absorptiokaistoista havaitsematta. Suuremmalla kanavien määrällä ja laajemmalla aallonpituusulottuvuudella varustetulla laitteistolla olisi todennäköisesti mahdollista tunnistaa kiviaineksia tarkemmin ja jakaa niitä pienempien ominaisuuksien muutosten mukaan eri ryhmiin. Ominaisuuksien muutoksilla tarkoitetaan tässä yhteydessä esimerkiksi muutoksia mineraalikoostumuksessa, raekoossa tai rapautumisasteessa.

Kiviaineksen lujuuteen ja rapautumisherkkyyteen vaikuttavat tekijät ovat spektriseen sormenjälkeen vaikuttavia ja näin ollen niiden tunnistamista ja luokittelua hyperspektridatan avulla voidaan pitää mahdollisena. Luokittelutulosten luotettavuutta lisäisi useampien aallonpituuskanavien käyttö tarjoamalla mahdollisuuksia hienovaraistenkin heijastusspektrin vaihtelevuuksien havainnoimisille. Koska tämä tekniikka eroaa selvästi avointen spektrikirjastojen mittaustekniikoista, eivät yleisesti saatavilla olevat referenssispektrit ole hyödynnettävissä. Laajempia geologisten kohteiden tutkimushankkeita varten olisi hyödyllistä tehdä omat spektrikirjastot ainakin yleisimpien kivien ja mineraalien osalta.

Käytännön applikaatioita varten mittausdatan esikäsittelyä olisi muokattava huomattavasti kevyempään suuntaan. Yksi mahdollisuus olisi käyttää superjatkumolaseria valonlähteenä aktiivisessa hyperspektrikuvantamisessa ja jättää lidarin antama paikkatieto käyttämättä. Varsinkin tuotannonaikaisessa laadunvalvonnassa ei spatiaalista tietoa välttämättä tarvittaisi. Laadukkaiden hyperspektrikuvien tuottaminen yhdistettynä tehokkaasiin luokittelualgoritmeihin ja kattaviin spektrikirjastoihin tarjoaisi varmasti hyvän työkalun geologisten kohteiden tutkimiseen.

Hyperspektrinen superjatkumo lidar osoittautui tässä tutkimuksessa erittäin mielenkiintoiseksi tutkimusvälineeksi ja sen ympärille on varmasti mahdollista kehittää monia käytännön sovellutuksia. Erilaiset kiviainesten laatuparametreihin ja muuntumisvyöhykkeiden tunnistamiseen liittyvät sovellutukset voisivat olla potentiaalisia kohteita. Käytetty tekniikka on vielä aika uutta ja se tulee varmasti ajan saatossa kehittymään tehokkaampaan ja tarkempaan suuntaan. Tämän tutkimuksen ja aiemman tiedon perusteella voidaan pitää todennäköisenä, että tämän kaltaisella laitteistolla on mahdollista tehdä kiviainesten luokittelua kohtuullisen suurella tarkkuudella.

## 8 Yhteenveto

Tämän diplomityön tarkoituksena oli selvittää, onko hyperspektrisen superjatkumo-lidarin avulla mahdollista luokitella kiviaineksia ennalta määrättyihin luokkiin. Tarkoituksena oli myös selvittää, olisiko hyperspektritiedon avulla mahdollista saada tietoa kiviainesten laatuun vaikuttavista tekijöistä, kuten raakoosta tai mineraalikoostumuksesta. Geologiselle aineistolle sopivien analysointimenetelmien selvittäminen ja vertailu nousivat myös tärkeään rooliin tulosten kannalta.

Tutkittavana oli seitsemän erilaista yleisessä käytössä olevaa rakennuskiveä, jotka olivat migmatiitti, gneissi, kvartsiitti ja neljä erilaista graniittia. Kivien heijastusspektri mitattiin Geodeettisen laitoksen kehittämällä hyperspektrisellä superjatkumo-lidarilla. Saadut spektritiedot jaettiin harjoitus- ja testialueisiin. Harjoitusalueilta poimittiin satunnaisotannalla mittauspisteitä, joiden avulla luotiin sekä referenssispektrit että harjoitussarjat. Referenssispektrejä käytettiin spektrikulman ja korrelaation mukaisissa luokitteluisissa sekä etäisyysluokittelussa. Harjoitussarjoja käytettiin ryhmittelyanalyysissä, kanonisessa erotteluanalyysissä ja pääkomponenttimuunnoksen avulla tehdyssä ryhmittelyanalyysissä.

Luokittelutulokset vaihtelivat kohtuullisesta erittäin hyvään. Heikoimmat tulokset tulivat spektrikulman ja korrelaation mukaisilla luokitteluilla. Tämä johtui harjoitussarjojen ja tuntemattomien näytteiden välisistä eroista, joille edellä mainitut menetelmät olivat erittäin herkkiä. Yksinkertainen euklidiseen etäisyyteen perustuva luokittelu osoittautui erittäin hyväksi menetelmäksi kyseiselle materiaalille. Sen antama luokittelun kokonaistarkkuus oli yli 96 prosenttia. Yhtä hyvät tulokset saatiin ryhmittelyanalyysillä, jossa yhdisteltiin hierarkista- ja k-keskiarvoanalyysiä. Parhaat tulokset tulivat ryhmien erottuvuutta maksimoivalla kanonisella erotteluanalyysillä. Sen avulla luokittelun kokonaistarkkuus oli yli 99 prosenttia. Kanonisen erotteluanalyysin todettiin olevan erittäin hyvä menetelmä toisiaan hyvin lähekkäisten luokkien erottamiseen. Myös pääkomponenttimuunnosta käytettiin dimensioiden vähentämisen vaikutusten tarkasteluun. Alkuperäiset kahdeksan muuttujaa esitettiin kolmen pääkomponentin avulla ja saadut ryhmittelyanalyysin tulokset jäivät alle yhden prosenttiyksikön päähän alkuperäisistä tuloksista.

Työn teoriaosuudessa käsiteltiin myös spektriseen sormenjälkeen vaikuttavia tekijöitä, ja niitä verrattiin kiviaineksen laatuun vaikuttaviin tekijöihin. Kiviaineksen laadun kannalta olennaiset raekoko ja mineraalikoostumus ovat molemmat selvästi spektriseen sormenjälkeen vaikuttavia tekijöitä. Tehtyjen havaintojen perusteella on oletettavaa, että kiviaineksen laatuparametreja voitaisiin estimoida hyperspektritiedon avulla. Myös olemassa oleva tutkimustieto viittaa siihen, että hyperspektritiedon ja kiviainesten laatuparametrien välillä on selkeä yhteys. Tämä olisi erittäin mielenkiintoinen jatkotutkimuskohde, jolla saattaisi olla merkitystä myös kiviainesteollisuudelle.

## Lähdeluettelo

ALHO, P., HYYPPÄ, H., HYYPPÄ, J., FLENER, C., KASVI, E., VAAJA, M., KUKKO, A., LOTSARI, E., HOHENTHAL, J., KURKELA, M., KAARTINEN, H. and HAGGREN, H., 2011. Uudet mittausmenetelmät jokiympäristön kartoituksessa. *The Photogrammetric Journal of Finland*, **22**(3), pp. 115-127.

ARNBOM, J., 2015-last update, Bergkvalitet - bergets lämplighet som ballastmaterial [Homepage of Sveriges Geologiska Undersökning (SGU)], [Online]. Available: [http://www.vattenfall.se/sv/file/Bergkvalitet\\_20140313.pdf\\_54553225.pdf](http://www.vattenfall.se/sv/file/Bergkvalitet_20140313.pdf_54553225.pdf) [1/25, 2016].

BALDRIDGE, A., HOOK, S., GROVE, C. and RIVERA, G., 2009. The ASTER spectral library version 2.0. *Remote Sensing of Environment*, **113**(4), pp. 711-715.

BARRETT, E.C., 2013. *Introduction to environmental remote sensing*. Routledge.

CAMPBELL, J. and WYNNE, R., 2011. *Introduction to remote sensing*. 5 edn. New York: The Guilford press.

CAMPS-VALLS, G., 2012. *Remote sensing image processing* /. San Rafael, Calif. (1537 Fourth Street, San Rafael, CA 94901 USA) : Morgan & Claypool.

CHESNER, W.H. and MCMILLAN, N.J., 2012. *Automated Laser Spectrographic Pattern Matching For Aggregate Identification*. 150. New Mexico: National Cooperative Highway Research Program.

CLARK, R.N., 25.6.1999, 1999-last update, Chapter 1: Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy, in *Manual of Remote Sensing, Volume 3, Remote Sensing for the Earth Sciences* [Homepage of U.S. Geological Survey], [Online]. Available: <http://speclab.cr.usgs.gov/PAPERS.refl-mrs/refl4.html> [15.6, 2015].

CLARK, R.N., GALLAGHER, A.J. and SWAYZE, G.A., 1990. Material absorption band depth mapping of imaging spectrometer data using a complete band shape least-squares fit with library reference spectra, *Proceedings of the Second Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Workshop 1990*, JPL Publication, pp. 4-5.

CLARK, R.N., SWAYZE, G.A., LIVO, K.E., KOKALY, R.F., SUTLEY, S.J., DALTON, J.B., MCDUGAL, R.R. and GENT, C.A., 2003. Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *Journal of Geophysical Research: Planets (1991–2012)*, **108**(E12),.

CLOUTIS, E.A., 1996. Review Article Hyperspectral geological remote sensing: evaluation of analytical techniques. *International Journal of Remote Sensing*, **17**(12), pp. 2215-2242.

DE CARVALHO, O.A. and MENESES, P.R., 2000. Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM), *Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18 2000*, JPL Publication Pasadena, CA.

- DOBIGEON, N., TOURNERET, J.-. and CHEIN-I CHANG, 2008. Semi-Supervised Linear Spectral Unmixing Using a Hierarchical Bayesian Model for Hyperspectral Imagery. *Signal Processing, IEEE Transactions on*, **56**(7), pp. 2684-2695.
- DUDLEY, J.M., GENTY, G. and COEN, S., 2006. Supercontinuum generation in photonic crystal fiber. *Reviews of Modern Physics*, **78**(4), pp. 1135.
- EKLUND, O., 2015-last update, Johdatus Etelä-Suomen kallioperään [Homepage of Suomen Kansallinen Geologian Komitea], [Online]. Available: <http://www.geologia.fi/index.php/2011-12-21-12-30-2011-12-21-12-39-11/2011-12-21-12-39-38/johdatus-etelae-suomen-kallioperaeen> [11/30, 2015].
- EVERITT, B., LANDAU, S., LEESE, M. and STAHL, D., 2011. *Cluster Analysis*. 5 edn. West Sussex, UK: John Wiley & Sons, Ltd.
- GAO, B., MONTES, M.J., DAVIS, C.O. and GOETZ, A.F., 2009. Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sensing of Environment*, **113**, pp. S17-S24.
- GUANG, Z. and MACLEAN, A.L., 2000. A comparison of canonical discriminant analysis and principal component analysis for spectral transformation. *PE&RS, Photogrammetric Engineering & Remote Sensing*, **66**(7), pp. 841-847.
- GUHA, A., RAO, D.A., RAVI, S., KUMAR, K.V. and RAO, E.N.D., 2012. Analysis of the potential of kimberlite rock spectra as spectral end member using samples from Narayanpet Kimberlite Field, Andhra Pradesh. *Current Science (00113891)*, **103**(9), pp. 1096-1104.
- HAKALA, T., SUOMALAINEN, J., KAASALAINEN, S. and CHEN, Y., 2012. Full waveform hyperspectral LiDAR for terrestrial laser scanning. *Optics express*, **20**(7), pp. 7119-7127.
- HASEGAWA, H., 2006. Evaluations of LIDAR reflectance amplitude sensitivity towards land cover conditions. *Bulletin of the Geographical Survey Institute*, **53**(6), pp. 43-50.
- HEIKKILÄ, J. and TÖRMÄ, M., 2003. *Hahmontunnistus : tilastollinen lähestymistapa*. 3. uud. p. edn. Espoo : Teknillinen korkeakoulu.
- HUGHES, G., 1968. On the mean accuracy of statistical pattern recognizers. *Information Theory, IEEE Transactions on*, **14**(1), pp. 55-63.
- HUNT, G.R., 1977. Spectral Signatures of Particulate Minerals in the Visible and Near Infrared. *Geophysics*, **42**(3), pp. 501-513.
- JÁNOS, T. and TÜNDE, T., 2008-last update, Geoinformatics [Homepage of Digitális Tankönyvtár], [Online]. Available: [http://www.tankonyvtar.hu/en/tartalom/tamop425/0032\\_terinformatika/ch05.html](http://www.tankonyvtar.hu/en/tartalom/tamop425/0032_terinformatika/ch05.html) [1/27, 2016].
- JOALA, V., 2006. *Laserkeilauksen perusteita ja mittauksen suunnittelu*. Espoo: Leica Nilomark Oy.

- JOLLIFFE, I.T., 2002. *Principal component analysis*. 2nd ed. edn. New York: Springer.
- KAURANNE, L.K., GARDEMEISTER, R., KORPELA, K. and MÄLKKI, E., 1972. *Rakennegologia II 304*. Otaniemi: Teknillisen korkeakoulun ylioppilaskunta.
- KUMAR, S., 2005. *Basics of remote sensing and GIS*. Firewall Media.
- LEHTINEN, M., NURMI, P. and RÄMÖ, T., 1998. *Suomen kallioperä: 3000 vuosisimiljöönaa*. Helsinki: Suomen Geologinen Seura ry.
- LEVERINGTON, D., 2009-last update, A Brief Overview of Reflectance Spectroscopy and Hyperspectral Imaging [Homepage of Texas Tech University, department of Geosciences], [Online]. Available: [http://www.webpages.ttu.edu/dleverin/hyperspectral/hyperspectral\\_imaging.html](http://www.webpages.ttu.edu/dleverin/hyperspectral/hyperspectral_imaging.html) [6/16, 2015].
- MANOLAKIS, D.G. and SHAW, G., 2002. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Processing Magazine*, **19**(1), pp. 29-43.
- MEER, F.V.D. and JONG, S.M.D., 2002. *Imaging spectrometry : basic principles and prospective application*. Boston : Kluwer Academic.
- NARENDRA, P.M. and FUKUNAGA, K., 1977. A Branch and Bound Algorithm for Feature Subset Selection. *IEEE Transactions on Computers*, **C-26**(9), pp. 917-922.
- NAVE, C.R., 2016-last update, Hyperphysics - Scattering from the Atmosphere [Homepage of Hyperphysics], [Online]. Available: <http://hyperphysics.phy-astr.gsu.edu/hbase/atmos/blusky.html#c4> [2/19, 2016].
- NEVALAINEN, O., HAKALA, T., SUOMALAINEN, J., MÄKIPÄÄ, R., PELTONIEMI, M., KROOKS, A. and KAASALAINEN, S., 2014. Fast and nondestructive method for leaf level chlorophyll estimation using hyperspectral LiDAR. *Agricultural and Forest Meteorology*, **198–199**, pp. 250-258.
- NIEMINEN, P. and JÄNISKANGAS, T., 1990. *Päällystekiviaineuksen rapautuminen ja sen arviointi. Asfalttipäällysteiden tutkimusohjelma ASTO 1987-1992*. 812. Espoo: VTT, tie- ja liikennelaboratorio.
- NISCHAN, M.L., NEWBURY, A.B., JOSEPH, R., IYENGAR, M.A., WILLARD, B.C., SWANSON, G.J., LIBBY, J., JOHNSON, B. and HSIAO-HUA, K.B., 2000. Active hyperspectral imaging, *International Symposium on Optical Science and Technology 2000*, International Society for Optics and Photonics, pp. 107-117.
- NISCHAN, M., JOSEPH, R., LIBBY, J. and KEREKES, J., 2003. Active spectral imaging.
- NURMIKOLU, A., 2004. *Murskatun kalliokiviaineuksen hienoneminen ja routivuus radan rakennekerroksissa*. Helsinki: Ratahallintokeskus.
- OLSEN, R.C., 2007. *Remote sensing from air and space* /. Bellingham, Wash. : SPIE Press.

ORIGINLAB, 2015-last update, Cluster Analysis (OriginPro) [Homepage of Origin lab], [Online]. Available: <http://www.originlab.com/doc/Origin-Help/Cluster-Analysis> [1/28, 2016].

REAGAN, J.A., MCCORMICK, M.P. and SPINHIRNE, J.D., 1989. Lidar sensing of aerosols and clouds in the troposphere and stratosphere. *Proceedings of the IEEE*, **77**(3), pp. 433-448.

SABINS JR, F.F., 1978. Remote sensing. *Remote sensing*, .

SCHOWENGERDT, R.A., 2007. *Remote sensing, models, and methods for image processing* /. 3rd ed. edn. Burlington, MA : Academic Press.

SERPICO, S.B., D'INCA, M., MELGANI, F. and MOSER, G., 2003. Comparison of feature reduction techniques for classification of hyperspectral remote sensing data, *International Symposium on Remote Sensing 2003*, International Society for Optics and Photonics, pp. 347-358.

SERPICO, S.B. and BRUZZONE, L., 2001. A new search algorithm for feature selection in hyperspectral remote sensing images. *Geoscience and Remote Sensing, IEEE Transactions on*, **39**(7), pp. 1360-1367.

SFS-EN 1097-2., 2010. *Kiviainesten mekaanisten ja fysikaalisten ominaisuuksien testaus. Osa 2: Iskunkestävyyden määrittämismenetelmät*. 2 edn. Helsinki: Suomen standardoimisliitto SFS.

SFS-EN 1097-9, 2014. *Kiviainesten mekaanisten ja fysikaalisten ominaisuuksien testaus. Osa 9: Nastarengaskulutuskestävyyden määrittäminen. Pohjoismainen testi (kuulamylymenetelmä)*. 2 edn. Helsinki: Suomen standardoimisliitto SFS.

SFS-EN 12407, 2009. *Natural stone test methods. Petrographic examination*. Helsinki: Suomen Standardoimisliitto SFS.

THEODORIDIS, S. and KOUTROUMBAS, K., 2009. *Pattern recognition* /. 4th ed. edn. Burlington, MA ; London : Academic Press.

VAN DER MEER, F., 2004. Analysis of spectral absorption features in hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, **5**(1), pp. 55-68.

VOSELMAN, G. and MAAS, H., 2011. *Airborne and Terrestrial Laser Scanning*. Whittles Publishing.

WANG, G. and WENG, Q., 2013. *Remote sensing of natural resources*. CRC Press.

WANG, L. and ZHAO, C., 2015. *Hyperspectral Image Processing*.

WIKLUND, S., 2008-last update, Multivariate Data Analysis for Omics [Homepage of Umetrics], [Online]. Available: [http://metabolomics.se/Courses/MVA/MVA%20in%20Omics\\_Handouts\\_Exercises\\_Solutions\\_Thu-Fri.pdf](http://metabolomics.se/Courses/MVA/MVA%20in%20Omics_Handouts_Exercises_Solutions_Thu-Fri.pdf) [1/25, 2016].

YUHAS, R.H., GOETZ, A.F. and BOARDMAN, J.W., 1992. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm, 1992.



## **Liiteluettelo**

Liite 1. Kuvat analyyseissä käytetyistä näytteistä. 7 sivua.



Kuva 1. Amadeus 3B



Kuva 2. Balmoral Red 3A





Kuva 3. Baltic Brown 12A





Kuva 4. Kuru Grey 8A



Kuva 5. Lappia Blue 12A





Kuva 6. Lieto Red 9A



Kuva 7. Nilsian kvartsiitti 15A